AL/HR-TR-1993-0056

AD-A267 019

ARMSTRONG LABORATORY

FLIGHT SIMULATOR-INDUCED SICKNESS AND VISUAL DISPLAYS EVALUATION

Harold D. Warner

University of Dayton Research Institute
300 College Park Avenue
Dayton, OH 45469-0110

Gary L. Serfoss, First Lieutenant, USAF
Todd M. Baruch

HUMAN RESOURCES DIRECTORATE
AIRCREW TRAINING RESEARCH DIVISION
558 First Street
Williams Air Force Base, AZ 85240-6457

David C. Hubbard

University of Dayton Research Institute
300 College Park Avenue
Dayton, OH 45469-0110

May 1993

Final Technical Report for Period October 1989 - January 1992

Approved for public release; distribution is unlimited.

AIR FORCE MATERIEL COMMAND
BROOKS AIR FORCE BASE, TEXAS

NOTICES

This technical report is published as received and has not been edited by the technical editing staff of the Armstrong Laboratory.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

Byron J. Pierce Dee H. Andrews

BYRON J. PIERCE
Project Scientist

DEE H. ANDREWS
Technical Director

Lynn A. Carroll

LYNN A. CARROLL, Colonel, USAF
Chief, Aircrew Training Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
May 1993

3. REPORT TYPE AND DATES COVERED
Final October 1989 - January 1992

4. TITLE AND SUBTITLE
Flight Simulator-Induced Sickness and Visual Displays Evaluation

5. FUNDING NUMBERS
C - F33615-90-C-0005
PE - 62205F
PR - 1123, 2743
TA - 32, 25
WU - 01, 17

6. AUTHOR(S)
Harold D. Warner, Todd M. Baruch
Gary L. Serfoss, David C. Hubbard

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Dayton Research Institute
300 College Park Avenue
Dayton, OH 45469-0110

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Armstrong Laboratory (AFMC)
Human Resources Directorate
Aircrew Training Research Division
558 First Street
Williams Air Force Base, AZ 85240-6457

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL/HR-TR-1993-0056

11. SUPPLEMENTARY NOTES
Armstrong Laboratory Technical Monitor: Dr. Byron J. Pierce, (602) 988-6561

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The objectives of this investigation were to compare the incidence and severity of simulator sickness associated with the use of two flight simulator visual systems and to evaluate the performance characteristics of the visual displays. The visual systems were the Display for Advanced Research and Training (DART) and the Limited Field-of-View Dome (LFOVD). Two groups of pilots served as subjects. One group consisted of active duty U.S. Air Force instructor pilots. The second group was comprised of pilots who were no longer flying military aircraft. The pilots performed a series of single-ship and formation flights, and an F-16 simulated fighter aircraft was used in conjunction with both visual systems. It was observed that both visual systems induced simulator sickness symptoms, and some of the pilots were forced to prematurely terminate the flights due to severe discomfort. For the pilots who were able to complete the flights, there was a significant increase in self reports of discomfort over time, an increase in simulator sickness symptomatology, and a decline in postural equilibrium. However, there were no differences in the incidence of simulator sickness between the two visual systems or pilot groups. The pilots recovered from the adverse effects of the simulation within 30 min following the flights. The means and medians of the pilots' ratings of the visual display characteristics indicated that the only problems encountered with the two visual systems were the resolution of the peripheral field and the vertical excursion limits of the high-resolution inset for the LFOVD. Flight performance data were also collected and analyzed. Significant differences in pilot performance were observed between the visual systems and pilot groups, and significant correlations were obtained between flight performance and the simulator sickness measures.

14. SUBJECT TERMS
Ataxia tests; Pilot discomfort ratings; Visual simulation systems; Flight performance; Postural disequilibrium; Flight simulators; Simulator sickness

15. NUMBER OF PAGES
114

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

CONTENTS

CONTENTS (continued)

List of Figures

List of Tables

List of Tables (concluded)

PREFACE

The present investigation was conducted at the Armstrong Laboratory's Aircrew Training Research Division (AL/HRA), Williams Air Force Base, AZ, to evaluate the incidence and severity of simulator-induced sickness associated with the use of two different flight simulator visual system technologies and to determine if any display design deficiencies existed that may have contributed to the malaise. This research effort was supported by the University of Dayton Research Institute (UDRI), Contract No. F33615-90-C-0005, in conjunction with Work Unit Nos. 2743-25-17, Flying Training Research Support, and 1123-32-01, Visual Display System Functional Requirements.

The authors wish to extend their deepest gratitude for the timely expertise provided in this investigation by the following individuals: Dr. Elizabeth Martin (AL/HRA), Debra Murray, Carlos Zavaleta, Jim Marx, Kathy Corral, Karen Hoffman (GEGS), Marge Keslin (UDRI), and the General Electric technicians who daily prepared the simulators for data collection.

FLIGHT SIMULATOR-INDUCED SICKNESS AND VISUAL DISPLAYS EVALUATION

INTRODUCTION

Simulator-induced sickness has been found to occur in conjunction with the use of various military flight trainers (e.g., Braithwaite & Braithwaite, 1990; Kennedy, Lilienthal, Berbaum, Baltzley, & McCauley, 1989; Parfitt & Chappelow, 1986). The percentage of users that experienced simulator sickness symptomatology ranged from 11% to 88%, depending upon the simulator (Casali & Frank, 1988). The symptoms of simulator sickness that were observed included: eyestrain, headaches, dizziness, sweating, drowsiness, and nausea. In some instances, aftereffects of the simulation have been reported in which the symptoms engendered during the simulation persisted for up to several days following the simulation, or the symptoms appeared only after the simulation ended as a delayed effect (Baltzley, Kennedy, Berbaum, Lilienthal, & Gower, 1989; Crowley, 1987; Ungs, 1988, 1989).

Reviews of the relevant technical literature (Casali, 1986; McCauley, 1984) indicate that a wide variety of variables may contribute to the incidence of simulator sickness. Examples of the more prominent variables are: temporal lags in the outputs of the visual and motion systems to pilot control inputs, frequency and acceleration of the motion system, field-of-view size and scene content of the visual system, cockpit environment factors such as temperature and humidity, duration and workload of the flight maneuvers, motion sickness susceptibility and experience level of the users, and simulator use characteristics such as freeze and reset. It is possible that these variables act in concert to produce the characteristic rates and severity of sickness symptomatology observed in relation to a simulator. That is, two or more of the variables may have to be operating simultaneously in order to induce adverse side effects (Casali, 1986).

The adverse effects of simulator-induced sickness have implications for the health and safety of the users as well as the training and research applications of the simulation. In terms of user safety, prolonged side effects or delayed symptoms could temporarily impair sensorimotor functions and interfere with such post-simulator activities as piloting an aircraft or driving a car. In the context of training, simulator sickness could, as has been pointed out (Kennedy, Berbaum, Allgood, Lane, Lilienthal, & Baltzley, 1988), promote distrust and apprehension of the simulation among users, which would compromise its training effectiveness. Furthermore, users could employ unrealistic responses to avoid or diminish adverse symptoms that could produce minimal, and possibly negative, transfer of training to the actual aircraft. The implications for research are that differences in user performance observed across the levels of the variables under investigation may be due to the different amounts of simulator sickness induced by the variables.

Rationale for the Present Evaluation

Two wide-angle visual simulation systems are currently in operation at the Aircrew Training Research Division of the Armstrong Laboratory (AL/HRA) to provide state-of-the-art research devices for the development of flight simulator visual display specifications. These visual systems are the Display for Advanced Research and Training (DART) and the Limited Field-of-View Dome (LFOVD). F-16 fighter cockpits are presently used in conjunction with both the DART and the LFOVD, and a wide range of flight operations can be simulated, such as: takeoffs and landings, formation flight, air refueling, low-altitude navigation, and air-to-ground weapon deliveries.

Much of the visual display research that will eventually be conducted using the DART and LFOVD will involve very low-level flight and vigorous maneuvering at high speeds. Because these conditions have been identified as potential contributors to simulator sickness, it is anticipated that some participants in the research may experience adverse reactions to the simulation. In order to determine if users will fall victim to simulator sickness under these demanding conditions, this investigation was conducted to assess the frequency and severity of simulator sickness associated with the use of both the DART and the LFOVD when similar tasks are performed.

It has been observed in previous research (Braithwaite & Braithwaite, 1990; Crowley, 1987; McGuinness, Bouwman, & Forbes, 1981) that in some instances older, more experienced pilots are more susceptible to simulator-induced sickness than younger, less experienced pilots. Due to the high probability that the ages and flight experience of the pilots participating in future DART and LFOVD research endeavors will vary extensively, two groups of pilots that differed with respect to age and flight experience were used in the present investigation. The two groups were compared to determine if pilot age and flight experience influenced the rates of simulator sickness occurrences in conjunction with the use of the DART and LFOVD, in order that the appropriate countermeasures could be adopted in subsequent investigations to alleviate the adverse effects of the simulation.

Purpose and Scope

The primary objectives of this investigation were to (a) identify and compare the frequency and severity of simulator sickness occurrences associated with the DART and LFOVD visual systems, and (b) assess the effects of pilot age and flight experience on the evocation of simulator sickness symptomatology. Secondary goals of the research were to (a) determine if any display deficiencies exist that may promote simulator sickness, and (b) evaluate pilot performance with the two simulation systems.

The participants were current and former military pilots. The pilots exercised complete control of the simulated aircraft, and they were requested to perform two very demanding tactical combat-like tasks. The tasks were a low-altitude, single-ship road reconnaissance task and a low-altitude formation flight task. Simulator sickness data were collected before, during, and after the simulator sessions. In addition, pilot opinions were solicited concerning the quality of the visual displays, and aircraft control data were obtained during the sessions.

METHODS

Subjects

Twenty-four pilots participated, 16 active duty U.S. Air Force T-37 and T-38 instructor pilots (IPs) and a group of 8 "mixed" pilots. The mixed group consisted of active duty military pilots and former military pilots who were not currently flying military aircraft. The average age and accumulated military flight hours of the IPs were 28.6 years and 1,228.4 hours, respectively. For the mixed pilots, the average age and military flight hours were 52.1 years and 3,879.6 hours. Contrast sensitivity was measured for each pilot using the Vistech Consultants, Inc., Vision Contrast Test System (VCTS), Model 6500. The tests indicated that the pilots had normal contrast sensitivity functions.

Visual Simulation Systems

The DART consists of a mosaic of eight pentagonal rear-projection windows that surround a simulator cockpit. An exterior view of the DART is shown in Figure 1 and an interior view in Figure 2. Computer-generated color imagery is projected onto the windows with commercial BARCO Electronic, BARCODATA 600 RGB, cathode ray tube (CRT) projectors. Each window is 55 inches (75 degrees) at the widest extent, and the eye-to-window distance is 37.5 inches. The total field of view is normally 300 degrees horizontal by 200 degrees vertical. The window directly above the pilot's head was not illuminated in this investigation. Consequently, the total field of view was 300 degrees horizontal by approximately 150 degrees vertical. A detailed description of the DART is provided by Thomas, Reining, and Kelly (1991).

A General Electric Advanced Visual Technology System (AVTS) computer-image generator is used to generate the visual imagery. The AVTS provides a total of ten video channels and the DART uses six of the channels. Since the DART is equipped with eight windows and each window requires one video channel, two windows are always

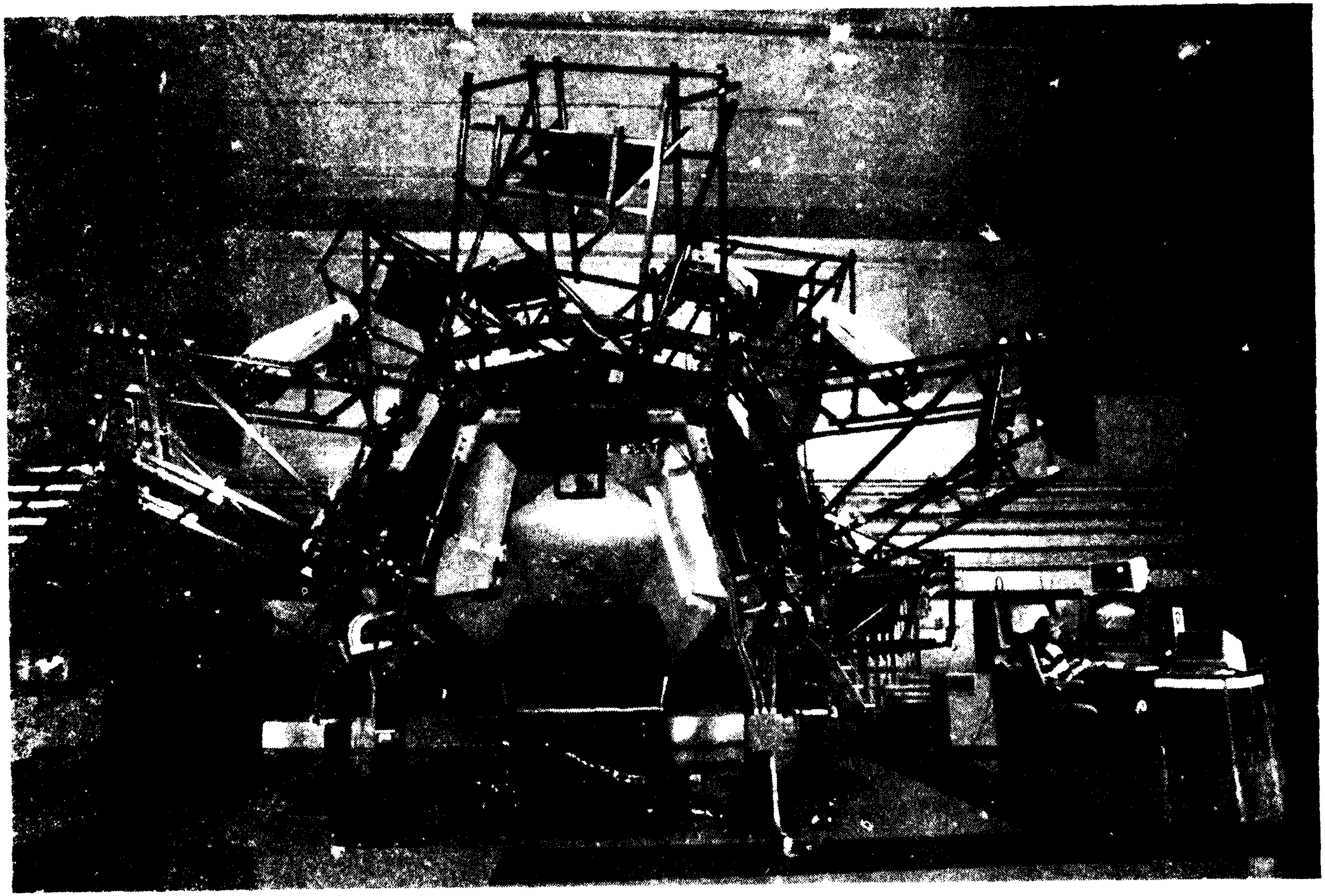

Figure 1
Exterior view of the DART

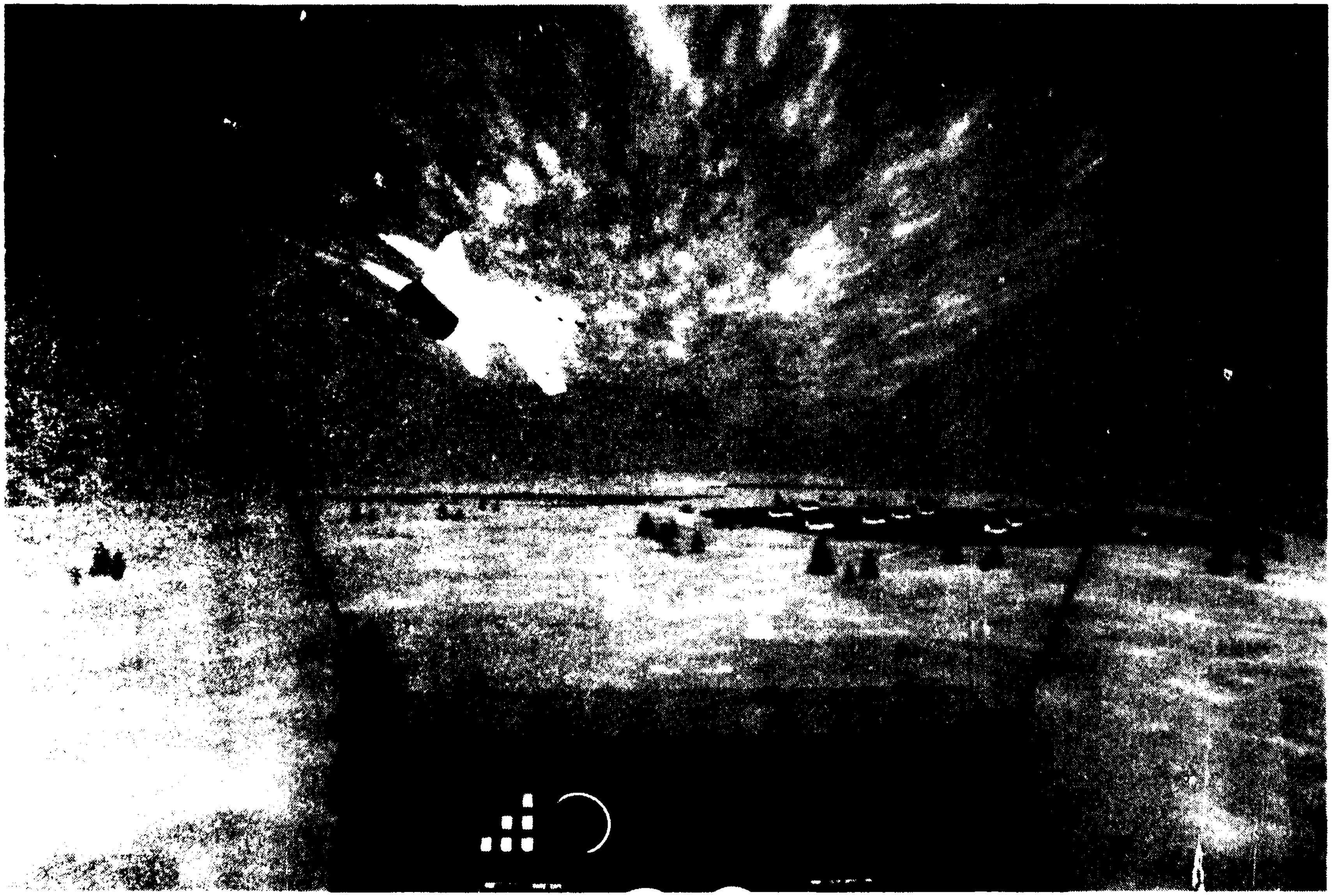

Figure 2
Interior View of the DART

blank. In the present investigation, the overhead window was blank as previously described, as well as the window on the side opposite the direction the pilot's head was turned. A Polhemus head-tracking system was used to measure head position for the window blanking.

A glass cockpit was enclosed within the DART and was configured to represent an F-16C fighter aircraft. An altimeter, attitude indicator, and airspeed indicator were presented on CRT displays on the cockpit instrument panel, and the cockpit was equipped with a side-arm pressure stick and a throttle with an afterburner. There were no rudder pedals. The trim switch and gun trigger on the control stick, as well as the speed brake switch on the throttle, were operational. F-16C head-up display (HUD) symbology was projected on the screen directly in front of the simulator cockpit. Two sets of indicator lights were installed in the cockpit below the glare shield. Each set contained three small indicator lights that were arranged in a vertical pattern. The top and bottom lights were amber, and the center light was green. The operation of these lights is described later in this report.

A simulator control station is located just outside the DART. The station provides a variety of controls and CRT monitors, which permits the operator and research personnel to control the study conditions and monitor the pilot's performance. A headset is provided in the simulator cockpit and a microphone and speaker are used at the control station to allow communications between the pilots and researchers.

The LFOVD visual simulation system (Fig. 3) employs an area-of-interest (AOI) display consisting of a movable, high-resolution inset that is surrounded by a wide-angle background display with lower resolution. The AOI and surrounding display are projected onto the interior surface of a 24-foot-diameter dome, and the eye-to-dome viewing distance is 12 feet. The field-of-view size of the AOI inset is 40 degrees horizontal by 30 degrees vertical, and the outside edge dimensions of the surrounding display are 140 degrees horizontal by 60 degrees vertical. Optical blending filters are used to provide a smooth transition between the high-resolution AOI inset and the lower resolution surrounding field. The size of the blend region is 5.0 degrees. The AOI is centered in the surrounding display and both are head slaved. The center point of the AOI can be rotated up to 90 degrees left and right, 40 degrees upward, and 22 degrees downward from a point directly in front of the simulator cockpit at eye level. A Polhemus head-tracking system was used to measure the pilot's head position and to position the visual scene on the dome surface.

The AVTS computer-image generator is used to provide the simulated visual scene. Two light-valve projectors are used to present the visual imagery on the dome, one for the AOI and one for the surrounding display. A comprehensive description of the output characteristics of the light valves is provided by Howard (1989).

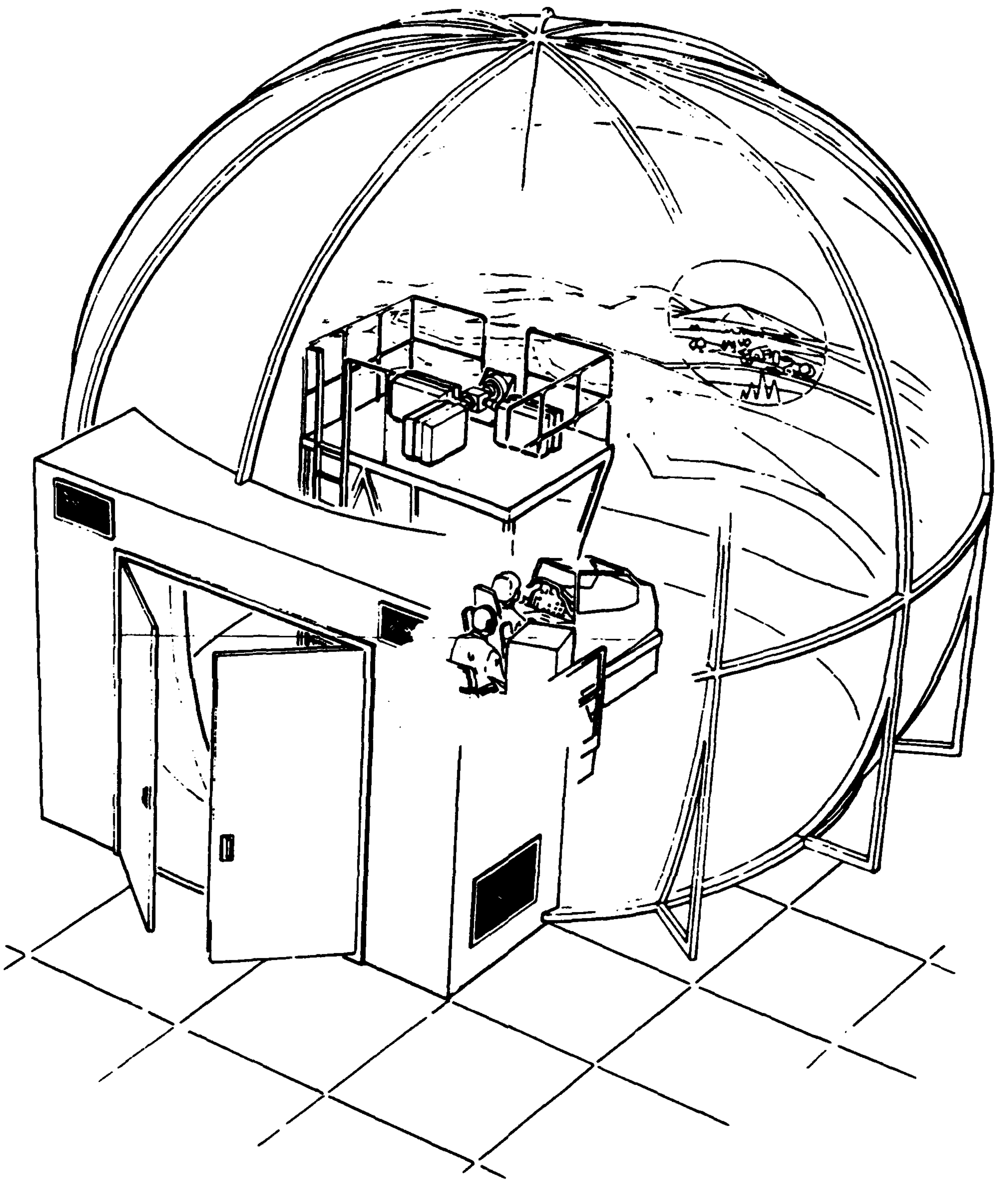

Figure 3
LFOVD Visual Simulation System

Optical lenses and filters are used to shape the AOI and surrounding field. The light-valve projectors and the AVTS servo-optical component, which contains the lenses and filters, are located in the dome above the simulator cockpit.

A fully operational F-16A simulator cockpit is enclosed in the dome. It provides actual F-16A aircraft controls, instrumentation, and a HUD. The controls consist of a side pressure stick with a trim switch and gun trigger, a throttle with afterburner and speed brakes, and rudder pedals. The instrumentation includes an altimeter, attitude indicator, and airspeed indicator. The HUD is located on the glare shield above the instrument panel. Two sets of indicator lights were also installed in the F-16A simulator cockpit, which were identical to the lights used in the cockpit associated with the DART.

The simulator control station is located in a room adjacent to the simulator. Controls and monitors are provided at the station allowing the operator and research personnel to control the study conditions and monitor the pilot's performance. A headset is provided in the simulator cockpit, and the control station is equipped with microphone and speaker for communications between the pilots and researchers.

Display Luminance, Contrast, and Modulation

The luminance and contrast characteristics of the DART and LFOVD are provided in Table 1. For the LFOVD, the measurements were taken in the AOI inset. It is evident in the table that the luminance and contrast of the DART were substantially higher than the LFOVD. The horizontal Modulation Transfer Functions (MTFs) associated with the DART and the LFOVD AOI are depicted in Figure 4. The figure shows that the modulation was greater for the DART at the lower spatial frequencies and greater for the LFOVD at higher spatial frequencies.

Table 1. Luminance and Contrast Characteristics of the DART and LFOVD

Display parameter	DART	LFOVD
Luminance (fL)		
Maximum	25.0	3.0
Minimum	0.5	0.3
Contrast ratio	50:1	10:1

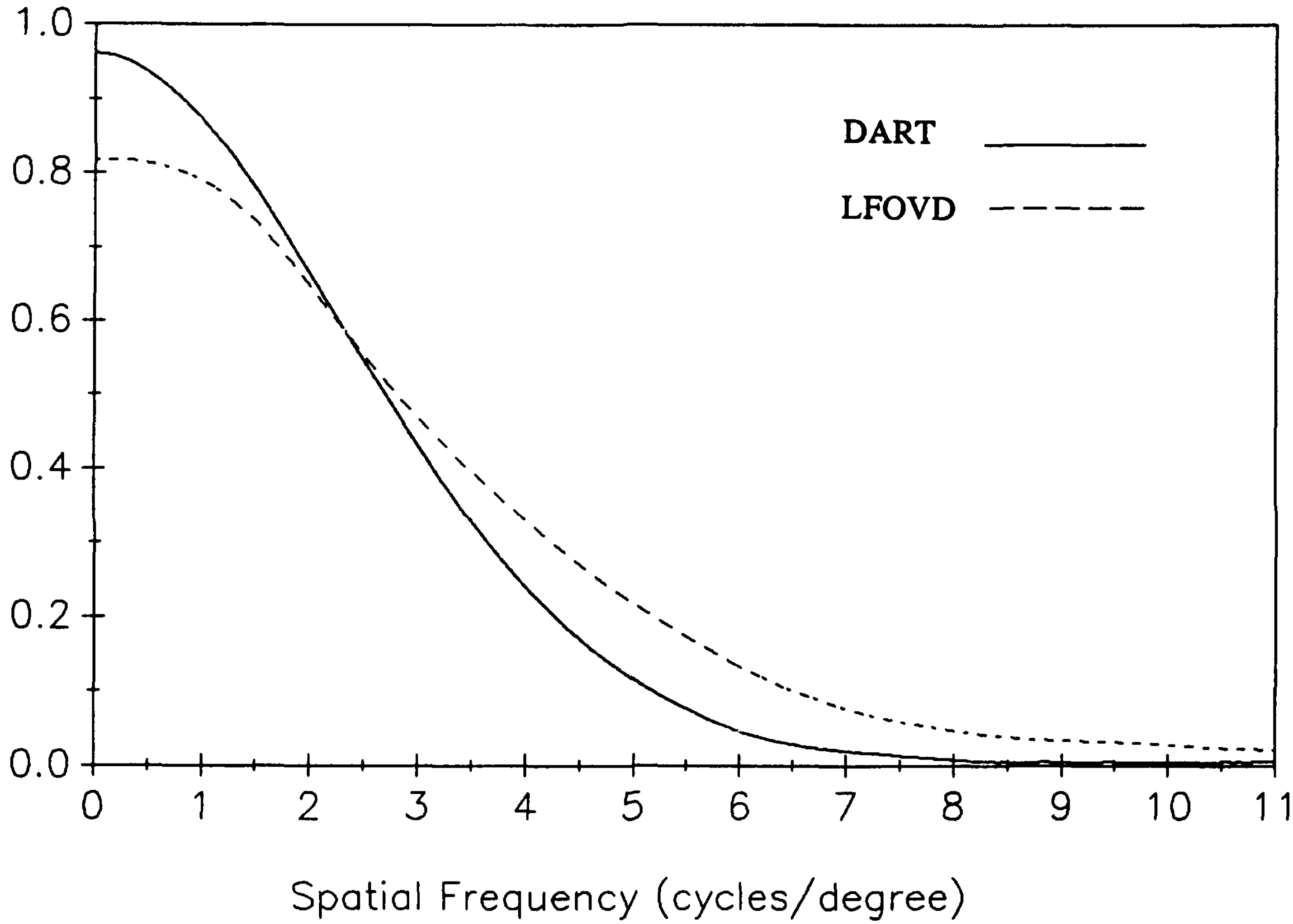

Figure 4
Horizontal MTFs for the DART and LFOVD AOI

Simulator Characteristics

An F-16A aerodynamics software package was employed in conjunction with the LFOVD F-16A simulator and an F-16C aerodynamics package was used with the DART F-16C simulator. A series of tests was devised to compare the performance characteristics of the LFOVD F-16A and the DART F-16C aerodynamics packages. The maneuvers that comprised the tests are identified in Table 2, and the time in seconds to accomplish the maneuvers is also provided. For these tests and during the investigation, the gross weight of the aircraft was 22,000 pounds, and the fuel quantity was frozen. The simulated aircraft carried no bombs, missiles, or external fuel tanks. Each maneuver was performed three times, and each value in the table represents the mean of the three trials. The "Time to climb at MAX power" is presented in the table for the actual F-16A and C models for comparative purposes. This value was obtained from the F-16A Flight Manual, T.O. 1F-16A-1, and the F-16C/D Flight Manual, T.O. 1F-16C-1. The flight manuals did not provide data for the other maneuvers that were accomplished in the performance tests. It may be observed in the table that the performance characteristics of the LFOVD F-16A were comparable to the DART F-16C.

There were two major differences, however, between the LFOVD F-16A and the DART F-16C simulators: control stick sensitivity and transport delay between control stick inputs and the output of the image generator. An actual F-16A control stick and force transducer were employed in the LFOVD F-16A simulator providing a stick sensitively in the simulator that was comparable to the actual aircraft. A commercially available force transducer, on the other hand, was used in conjunction with the DART F-16C simulator. This transducer was far more sensitive than the actual aircraft transducer, meaning that it took less stick pressure to pitch and bank the simulated aircraft to the same extent as the LFOVD F-16A simulator. Another way of stating this is that for the same amount of stick pressure, the DART F-16C simulator pitched and banked at a higher rate than the LFOVD F-16A simulator.

The transport delay was also greater for the DART F-16C simulator than for the LFOVD F-16A simulator. Leinenwever and Moran (1992) compared the transport delays of the two simulators (along with a third flight trainer), prior to optimizing the hardware and software, using a test methodology in which the times between a discrete stick input signal and the outputs of the AVTS image generator and the cockpit attitude director indicator (ADI) were measured. In these tests, a pitch change input signal was used and the output was a $\pm$ 80 degree change in pitch. Both best-case and worst-case sampling conditions were measured. In the best-case condition, the input signal occurs immediately before the stick state is sampled, which eliminates one computational cycle in the transport delay measurement. In the worst-case condition, the

Table 2. Comparison of Performance Characteristics Between the Actual F-16A/C, LFOVD F-16A, and DART F-16C

Maneuver	Actual F-16A/C	LFOVD F-16A	DART F-16C
Time to climb at MAX power from 2,500 ft. MSL to 40,000 ft. MSL starting at 550 KIAS	75.00	76.00	70.33
MAX power acceleration from 350 KIAS to 550 KIAS at 1,000 ft. MSL	N.A.	14.67	13.67
MIL power acceleration from 350 KIAS to 550 KIAS at 1,000 ft. MSL	N.A.	30.00	29.00
IDLE power deceleration from 550 KIAS to 350 KIAS at 1,000 ft. MSL with speed brakes deployed	N.A.	17.00	16.33
720 degree roll with a maximum roll rate starting at 250 KIAS and at 7,000 ft. MSL	N.A.	4.00	4.00
720 degree roll with a maximum roll rate starting at 550 KIAS and at 7,000 ft. MSL	N.A.	4.00	4.00

Notes. 1. Values in seconds.
2. N.A. = not available in flight manuals.
3. KIAS = knots indicated airspeed.
4. MSL = mean sea level.

input signal occurs immediately after the stick state is sampled, which adds a full computational cycle to the transport delay measurement. The transport delays of the LFOVD F-16A and DART F-16C simulators observed by Leinenwever and Moran are presented in Table 3.

Visual Environment

The simulated visual environment consisted of a continuous canyon with a narrow floor and high, sloping walls. The pilot's view of the canyon from the cockpit is shown in Figure 2, and an overhead view of the entire canyon is depicted in the line drawing in Figure 5. A green, forest texture pattern was placed on the canyon walls, which rose to a maximum of 4,500 feet above sea level. A blue-gray cloud pattern was used for the sky, and the

Table 3. Transport Delays of the DART F-16C and LFOVD F-16A Simulators

Output	DART F-16C Best	DART F-16C Worst	LFOVD F-16A Best	LFOVD F-16A Worst
Image generator	183.6	211.6	97.7	127.5
ADI	155.9	222.7	61.8	76.7

Note. Values in milliseconds.

clouds moved as though being pushed by the wind. A light-green texture pattern was used for the canyon floor, which varied in elevation from 10 feet up to 600 feet above sea level over the entire length of the canyon. There was a tan-colored road on the floor of the canyon. Numerous lakes, farm fields, and two-dimensional textured shapes representing residential areas were also placed in the canyon. Additionally, there was a wide variety of three-dimensional objects, including a large number of trees, convoys of military vehicles on the road, houses, farm buildings, churches with steeples, a football stadium, a drive-in movie theater, and a lumber mill.

The same visual environment was used with both the DART and LFOVD. Because the display luminance of the LFOVD was substantially lower than the luminance of the DART, it was not possible to equate the luminance levels of the visual environment in the LFOVD with the luminance levels in the DART. Alternatively, an approach was adopted in which the contrast levels between objects or textured surfaces and their backgrounds in the DART were matched with the corresponding contrast levels in the LFOVD in the following manner. First, the image generator color tables for the LFOVD, which determine the colors and luminance levels in the visual scene, were adjusted to provide a realistic portrayal of the canyon environment. Luminance measurements were then obtained for a wide range of objects and textured surfaces in the environment, and the contrast levels were calculated using the Michelson equation for contrast as follows:

$$\text{Contrast} = \frac{A - B}{A + B}$$

where A is the luminance of one surface and B is the luminance of another surface. Next, the visual environment was displayed in the DART, and the luminance levels of the same objects and textured surfaces were measured in the display window directly in front of the simulator cockpit. The color tables for the DART were then adjusted to obtain the required subjective color matches and

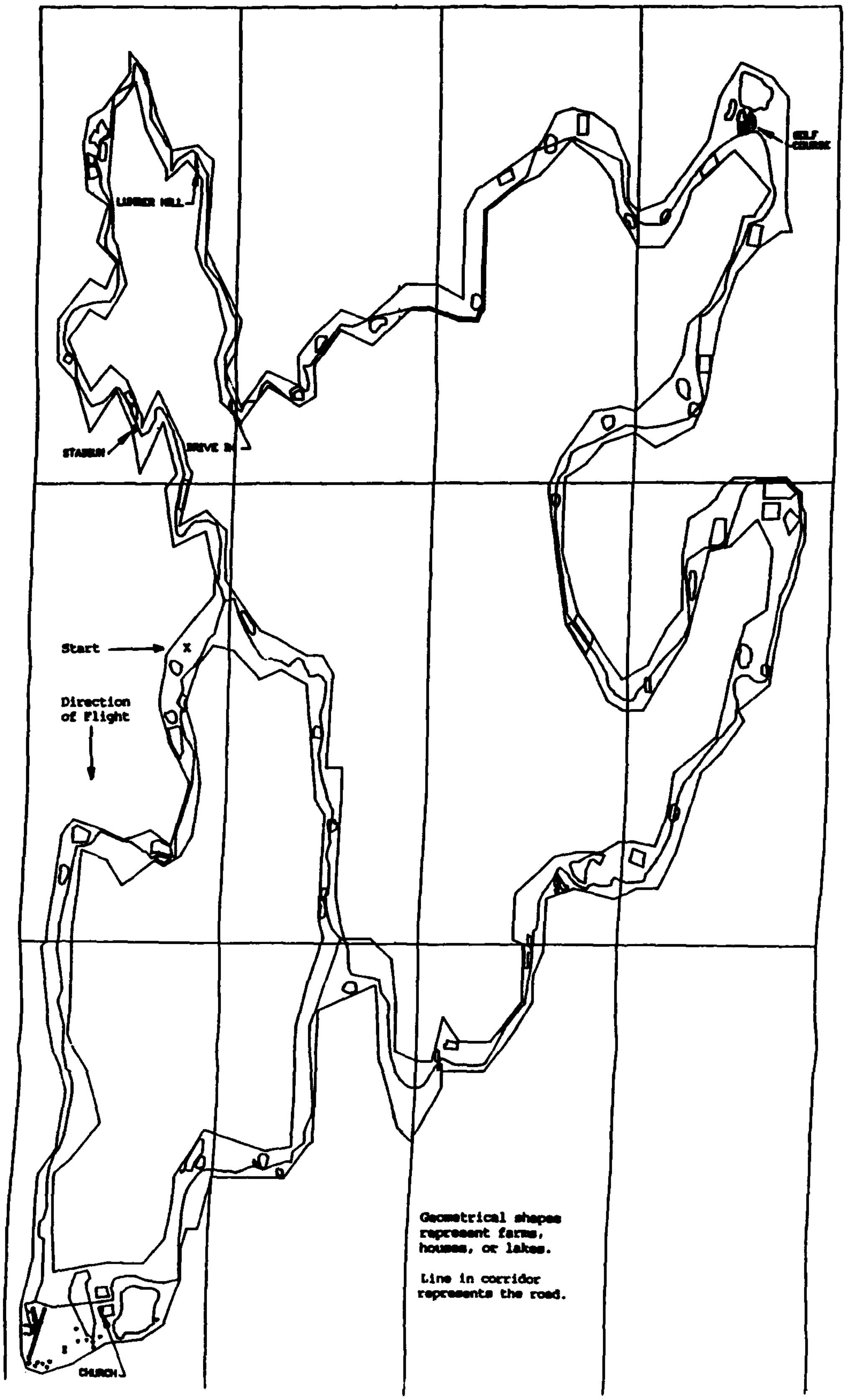

Figure 5
Overhead Representation of the Canyon Visual Database

measured luminance contrast matches. Since separate color tables were used for each of the DART display windows, an automated window calibration procedure developed by Dr. Celeste Howard at the University of Dayton Research Institute was implemented to equate the colors and luminance levels in the display windows.

Performance Task

The task consisted of: (a) a 5-min simulator familiarization flight, (b) a 5-min practice formation flight, (c) a 20-min single-ship flight, and (d) a 20-min formation flight. In the familiarization and single-ship flights, the pilots were instructed to bracket their airspeed between 425 and 475 knots and stay below 500 feet AGL (above ground level). To assist the pilots in adhering to these requirements, the indicator lights in the cockpit were used to inform the pilots whether they were within or outside of the allowable airspeed and altitude limits. When the airspeed was within the allowable range, the middle, green indicator on the left-hand set of lights was illuminated and the upper and lower amber lights were off. If the speed was too fast, the indicator lights illuminated as follows: the center and upper lights flashed between 476 and 500 knots, between 501 and 550 knots, the upper light only flashed, and the upper light was steady on above 551 knots. If the speed was too slow: the center and lower lights flashed between 424 and 400 knots, between 399 and 350 knots, the lower light only flashed, and the lower light was steady on below 349 knots. The center, green and top, amber indicators on the right-hand set of lights were used to provide altitude feedback. When the aircraft was 500 feet AGL or below, the center light was illuminated and the top light was off. If the aircraft exceeded 500 feet, the top light flashed and the center light was off. In addition to the airspeed and altitude requirements, the pilots were requested to follow the road in the canyon for the single-ship flight and verbally signal the presence of the vehicles on the road. They were asked to follow the road to induce more vigorous maneuvering. The pilot's verbal signal was required to determine whether the pilots were following and visually scanning the road; their responses were not recorded.

For the practice and 20-min formation flights, a highly detailed MiG-29 served as the lead aircraft. The lead aircraft was prerecorded, and the same recording was used in each formation flight. The first 5 min of the prerecorded flight were used for formation practice. The lead aircraft flew very rigorous low-level maneuvers, which entailed variations in altitude and frequent turns.

The pilots were instructed to fly in formation with a slant range separation of 2,000 feet or less. The indicator lights on the right-hand side in the cockpit were used to provide slant range feedback. The center, green light was continuously illuminated when the aircraft was within 2,000 feet of the lead aircraft. If the pilot exceeded the allowable limit, the center, green and bottom, amber lights flashed when the slant range was between 2,001 and 3,000 feet, the bottom only light flashed when the range was between 3,001 and 5,000 feet, and the bottom light was steady on if the range was greater than 5,000 feet. The airspeed and altitude feedback indicator lights were not illuminated during the formation flights.

If the aircraft contacted any objects or the terrain or penetrated the clouds, the visual scene flashed red and was masked with a uniform gray color. The aircraft was then automatically repositioned in the canyon, the mask was removed, and the flight continued. This crash and cloud penetration reset feature was used in both the single-ship and formation flights. A control function was also provided at the simulator control stations that allowed the researchers to re-form the pilot with the lead aircraft in the formation flights if the pilot lost sight of the lead aircraft and was unable to reestablish the formation. If the pilot had to be re-formed with the lead, the display was masked while the formation was reestablished.

The pilots were permitted to fire the simulated aircraft guns during the flights. Tracer rounds were displayed when the gun was fired, but there was no smoke nor damage feedback if an object was hit. The pilot could also fire at the lead aircraft, but it could not be shot down.

Flight Performance Data

Raw flight performance data were collected during the single-ship and formation flights, then the data were scored off-line for use in the statistical analyses. The flight performance data obtained, the units of measurement, and the scores produced for the analyses are shown in Table 4.

The data were collected at a rate of two samples per s. Terrain crash frequency and cloud penetration frequency were also recorded during the single-ship and formation flights, and the frequency that the pilots were manually re-formed with the lead aircraft was obtained in the formation flights.

Simulator Sickness Assessment

A battery of questionnaires and ataxia (i.e., postural equilibrium) tests that have been applied in other simulator sickness evaluations (Kennedy et al., 1988; McCauley, Hettinger, Sharkey, & Sinacori, 1990) was administered to the pilot participants before and after the simulator sessions. The questionnaires consisted of a motion history questionnaire (MHQ)

Table 4. Flight Performance Measures

Flight	Data	Units	Scores
Single-ship and formation	g-loading	g's	mean standard deviation
	altitude	feet AGL	mean standard deviation
	airspeed	knots	mean standard deviation
	pitch angle	degrees	absolute mean standard deviation
	bank angle	degrees	absolute mean standard deviation
	pitch rate	degrees/ second	absolute mean standard deviation
	roll rate	degrees/ second	absolute mean standard deviation
	yaw rate	degrees/ second	absolute mean standard deviation
Formation only	slant range	feet	mean standard deviation

and a simulator sickness questionnaire (SSQ); the ataxia tests were: Stand on Leg Eyes Closed (SOLEC) and Walk on Floor Eyes Closed (WOFEC). The MHQ is designed to identify an individual's prior motion sickness occurrences and susceptibility to motion sickness. The MHQ can be scored and the scores then correlated with simulator sickness incidents to determine how well the MHQ predicts simulator sickness occurrences in association with the use of the DART and LFOVD. Table 5 contains the MHQ that was administered to the pilots.

The SSQ used in the investigation is presented in Table 6. The pilots were to indicate all the symptoms they were experiencing at the time the SSQ was administered. The ataxia tests were employed as a means of measuring the pilots' postural equilibrium, which can be disrupted as a function of exposure to a simulator. In the SOLEC test, the pilots were required to stand on one leg with their eyes closed, and the time the pilots were able

Table 5. Motion History Questionnaire for Simulator Sickness (from Kennedy et al., 1988)

1. Have you ever been motion sick other than aboard ships or in aircraft?

No _____ Yes_____

2. Listed below are a number of situations in which some people have reported motion sickness symptoms. In the space provided, check any SYMPTOM(S) you may have experienced at any time, past or present.

SITUATIONS	SYMPTOMS												
	VOMITED	NAUSEA	STOMACH AWARENESS*	INCREASED SALIVATION	DIZZINESS	DROWSINESS	SWEATING	PALLOR	VERTIGO	AWARENESS OF BREATHING	HEADACHE	OTHER SYMPTOMS	NONE
AIRCRAFT													
FLIGHT SIMULATOR													
ROLLER COASTER													
MERRY-GO-ROUND													
OTHER CARNIVAL DEVICES													
AUTOMOBILES													
LONG TRAIN OR BUS TRIPS													
SWINGS													
HAMMOCKS													
GYMNASTIC APPARATUS													
ROLLER/ICE SKATING													
ELEVATORS													
CINERAMA OR WIDE-SCREEN MOVIES													
MOTORCYCLES													

* Stomach Awareness refers to a feeling of discomfort that is preliminary to nausea.

Table 6. Simulator Sickness Questionnaire (SSQ) (adapted from Kennedy et al., 1989)

Instructions: For each symptom, circle the rating that applies to you RIGHT NOW.

	SYMPTOM	RATING			
1.	General discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Boredom	None	Slight	Moderate	Severe
4.	Drowsiness	None	Slight	Moderate	Severe
5.	Headache	None	Slight	Moderate	Severe
6.	Eye strain	None	Slight	Moderate	Severe
7.	Difficulty focusing	None	Slight	Moderate	Severe
8a.	Salivation increased	None	Slight	Moderate	Severe
8b.	Salivation decreased	None	Slight	Moderate	Severe
9.	Sweating[a]	None	Slight	Moderate	Severe
10.	Nausea	None	Slight	Moderate	Severe
11.	Difficulty concentrating	None	Slight	Moderate	Severe
12.	Mental depression	No	Yes		
13.	"Fullness of the head"	No	Yes		
14.	Blurred vision	No	Yes		
15.	Dizziness	No	Yes		
16.	Vertigo	No	Yes		
17.	Visual flashbacks[b]	No	Yes		
18.	Faintness	No	Yes		
19.	Aware of breathing	No	Yes		
20.	Stomach awareness[c]	No	Yes		
21.	Loss of appetite	No	Yes		
22.	Increased appetite	No	Yes		
23	Desire to move bowels	No	Yes		
24.	Confusion	No	Yes		
25.	Burping	No	Yes		
26.	Vomiting	No	Yes	No. of times_____	
27.	Other: Please describe_______________				

a "Cold sweating" due to discomfort, not due to physical exertion.
b "Visual flashback" is a visual illusion of movement or false sensations similar to aircraft dynamics when NOT in the simulator or aircraft.
c "Stomach awareness" is usually used to indicate a feeling of discomfort just short of nausea.

to maintain this posture was recorded up to a maximum of 30 s. For the WOFEC test, the pilots walked heel-to-toe on a straight line with their eyes closed, and the number of steps they accomplished before stepping off the line were recorded up to a maximum of 12 steps. Comprehensive descriptions of these postural equilibrium tests can be found in Fregly, Graybiel, and Smith (1972), Hamilton, Kantor, and Magee (1989), and Thomley, Kennedy, and Bittner (1986).

Along with the questionnaires and ataxia tests that were administered, discomfort ratings were obtained from the pilots during the simulator flights. A 7-point scale was adopted, where 1 indicated "normal, symptom free" and 7 indicated "severe discomfort."

Visual Display Evaluation

The pilots were also asked a series of display-related questions during the simulator flights, and they were administered a comprehensive display evaluation questionnaire after the conclusion of the flights in both simulators. The questionnaire was comprised of a list of display characteristics, and the pilots rated the acceptability of each characteristic. A 5-point scale, which was extracted from Meister (1986), was used:

___ Very Acceptable
___ Acceptable
___ Borderline
___ Unacceptable
___ Very Unacceptable

The pilots were required to identify the deficiencies they observed when they rated the display characteristic "borderline" or below. The display-related questions asked during the flights and the display characteristics addressed in the display evaluation questionnaires are provided in Appendixes A and B along with the pilot responses.

Procedure

Each pilot participated in two simulator sessions, one with the LFOVD and one with the DART. The order in which the sessions were accomplished was counterbalanced across the two groups of pilots, and the sessions were at least two weeks apart. Within each session, the pilots first performed the 5-min simulator familiarization flight, then the 5-min practice formation flight, and then either the 20-min single-ship flight followed by the 20-min formation flight or the formation flight followed by single-ship flight. The order in which the 20-min single-ship and formation flights were presented in the sessions was counterbalanced.

The MHQ was administered to each of the pilots prior to their first simulator session. The ataxia (SOLEC and WOFEC) tests and the SSQ were administered three different times per simulator session: just before the participants entered the simulator, immediately after the session ended, and again 30 min following the session. The ataxia tests were applied four times at each of the three time periods: one practice test and three repetitions. The mean of the three repetitions was used in the analysis of the ataxia data.

Ten pilot discomfort ratings were requested in each session. The ratings were obtained at the end of the simulator familiarization flight, at the end of the practice formation flight, and at 5-min intervals in the single-ship and formation flights. The procedures adopted in the present investigation concerning the administration of the questionnaires and ataxia tests and the collection of the pilot discomfort ratings were based on the approach used by McCauley et al. (1990).

The display-related questions were asked during the 20-min single-ship and formation flights, and a verbal task difficulty rating was obtained at the end of each flight segment. One display question was presented at 5 min, one at 10 min, and one at 15 min within each flight in the order they are listed in Appendix A. The questions were asked immediately after the pilot discomfort ratings were recorded by the experimenter. For the task difficulty ratings, the pilots were asked to judge difficulty on a 7-point scale, where 1 was "very easy" and 7 was "very difficult." The display evaluation questionnaires were administered at the conclusion of the simulator sessions.

Standardized instructions were provided prior to the start of each simulator session after the pilots were seated in the cockpit. In the instructions, the operation of the simulator, the visual simulation system, the task requirements, and the procedures were described. The pilots were also informed that they could terminate the session at any time.

The luminance levels associated with the LFOVD were calibrated each day before the first session commenced. This was required because the light-valve luminances were frequently changed to support other research endeavors and because of the deteriorization of the arc lamps. The DART projector luminances were very stable and did not require recalibration.

RESULTS

Session Terminations

Due to severe discomfort, one IP and two of the pilots in the "mixed" group were unable to complete the simulator sessions with either the LFOVD or the DART visual systems. Also, one of the pilots in the mixed group had to terminate the session when the DART was used, but was able to complete the session with the LFOVD. Following this latter session, however, the pilot exhibited severe sweating and pallor, which are prominent overt symptoms of simulator sickness. The times at which the sessions were terminated by the pilots are shown in Table 7. The times indicated in the table represent the total accumulated flight time from the start of the initial simulator familiarization flight. Overall, 16.67% of the pilots terminated the sessions with the DART because of severe discomfort, and 12.5% were unable to complete the flights with the LFOVD.

Table 7. DART and LFOVD Session Terminations

Group	Pilot ID No.	DART	LFOVD
IPs	10	16:16	38:55
Mixed	1	24:52	30:00
	4	12:17	16:56
	7	26:42	(Completed)

Note. Times are in min:sec.

Pilot Discomfort Ratings

The individual discomfort ratings obtained from each of the pilots unable to complete the simulator sessions using the DART and LFOVD are presented in Figures 6 and 7, respectively. The last rating for each pilot represents the final rating collected before the pilot terminated the session; it does not necessarily reflect the pilot's discomfort when the pilot actually requested cessation of the simulation. Due to the peculiar trend in the ratings for IP number 10 with the LFOVD visual system (Fig. 7), a brief explanation is warranted. This pilot experienced increasing discomfort between 10 and 30 min, which was during the 20-min formation flight task. When the 20-min single-ship flight task was subsequently begun, the pilot reported feeling much better. At 8 min and 55 s into the single-ship flight task, however, the pilot experienced a sudden onset of symptoms and asked that the flight be stopped.

To determine whether the remaining pilots experienced a significant increase in discomfort, the discomfort ratings of the

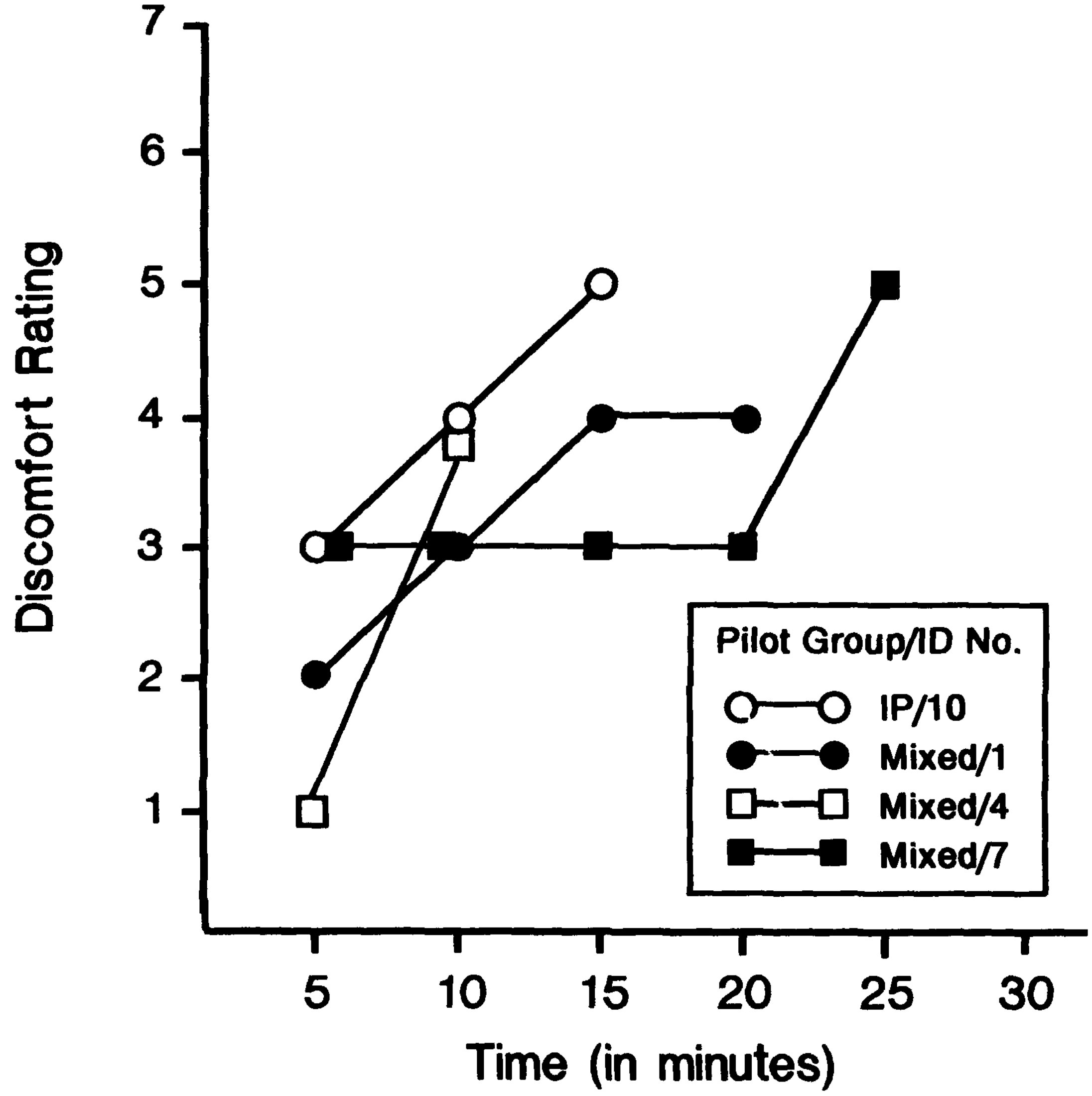

Figure 6
Discomfort Ratings of the Pilots Who Were Unable to Complete the Flights with the DART

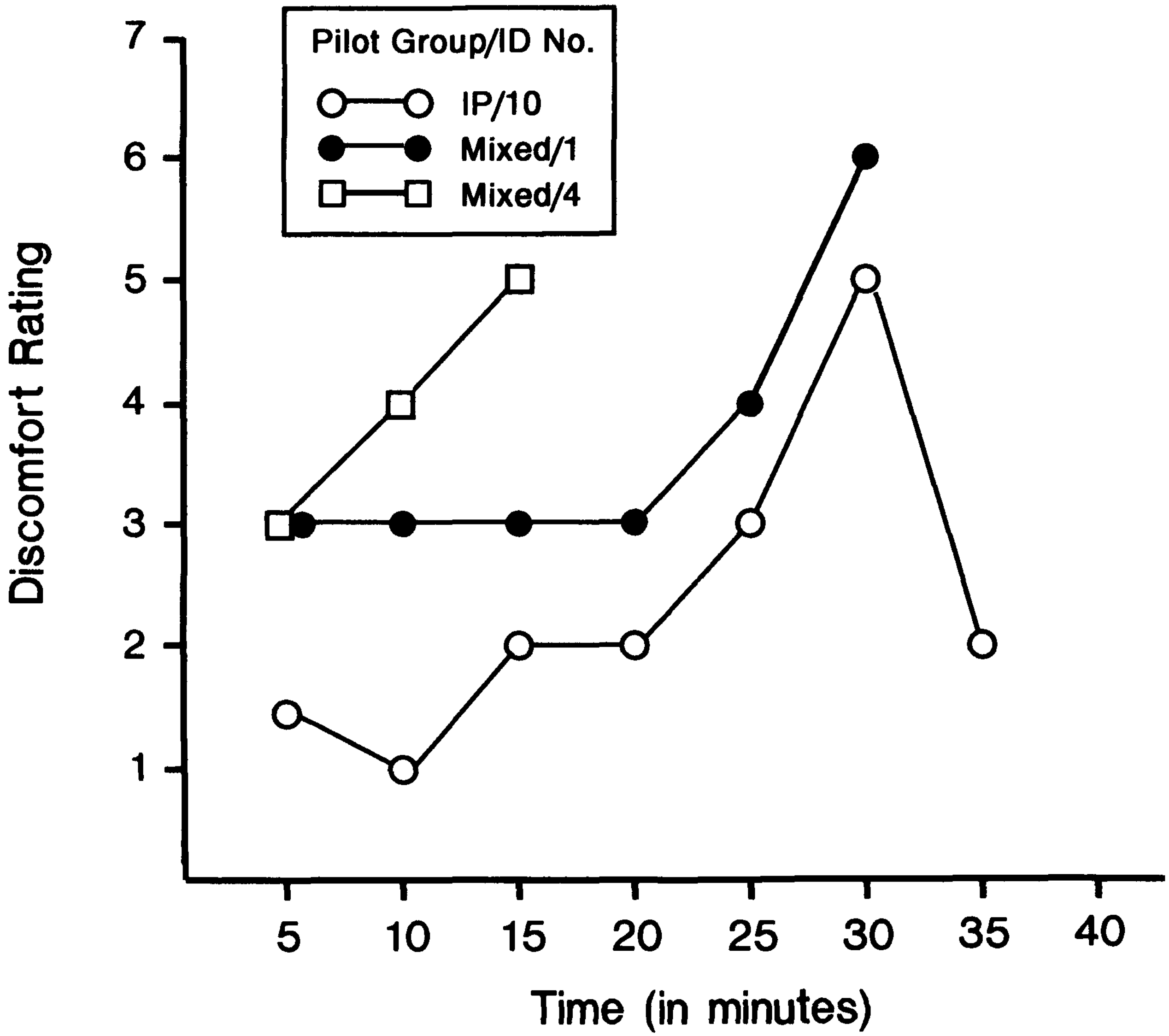

Figure 7
Discomfort Ratings of the Pilots Who Were Unable to Complete the Flights with the LFOVD

pilots who completed the sessions were subjected to a repeated measures analysis of variance (ANOVA) in which pilot group was the between-subjects factor and the within-subjects factors were visual system and flight time. The analysis indicated that the main effect of time was statistically significant, $F(9,162) = 2.97$, $p < 0.01$. The distribution of the mean discomfort ratings is presented in Figure 8. The ratings increased on the average up to the 30-min mark where the task was changed. The ratings declined at this point, but then increased to an even higher mean level. The visual system and pilot group main effects were not significant nor were the interactions.

Simulator Sickness Questionnaire (SSQ)

The severity level of the symptoms that increased between the pre and post 1 SSQs for the pilots who were unable to complete the sessions are shown in Tables 8 and 9 in relation to the DART and LFOVD, respectively. Pilot number 1 in the mixed group was unable to complete the simulator session with either the DART or the LFOVD, but was inadvertently not administered the SSQ after the session with the DART was terminated. It will be recalled that pilot number 7 in the mixed group was able to complete the session with the LFOVD but not the DART. The post 1 SSQ was administered to the pilots that terminated the sessions when they felt well enough to complete it. Because the pilots had recovered from the adverse effects of the simulation to some degree, the severity of the symptoms identified in Tables 8 and 9 do not necessarily reflect the actual extent of the malaise at the moment the sessions were halted. The post 2 SSQ was not given to the pilots who were unable to complete the sessions because the intent of the questionnaire, which was to determine how much they had recovered within 30 min following the sessions, had been invalidated by the necessity for waiting until the pilots felt better to complete the post 1 symptom checklist.

The SSQs of the pilots who were able to complete the simulator sessions were scored and the scores were subjected to an ANOVA. The SSQs were scored using a computerized algorithm provided by Monterey Technologies, Inc., which was based upon the SSQA scoring method developed by Lane and Kennedy (1988). This algorithm converted the pilot's symptom checklist responses to a single, composite score with a minimum value of 100.0, which reflected the absence of simulator sickness symptomatology. The algorithm also partitioned the responses into separate nausea, visuomotor, and disorientation scores, but these scores were not used in the present investigation. The algorithm retained 15 of the symptoms from the checklist and omitted 12 checklist symptoms.

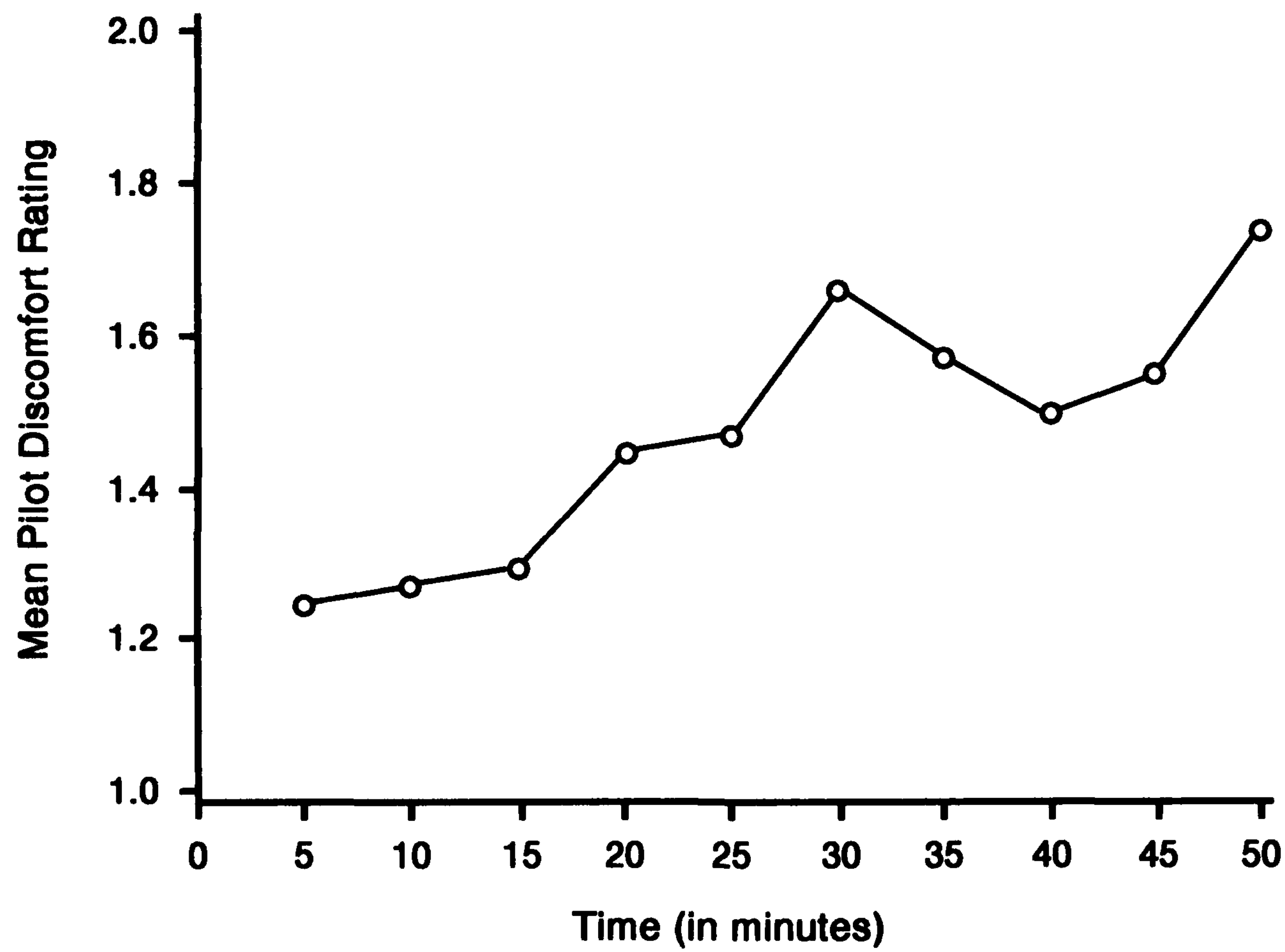

Figure 8
Mean Discomfort Ratings of the Pilots Who Were Able to Complete the Flights

Table 8. Simulator Sickness Symptomatology of the Pilots Who Were Unable to Continue the Sessions Using the DART

Symptom	Pilot group/ID no. IP/10	Mixed/4	Mixed/7
General discomfort	Moderate	Moderate	Severe
Salivation increased	Slight	-	Slight
Sweating	Moderate	Moderate	Severe
Nausea	Slight	Slight	Severe
Difficulty concentrating	Slight	-	Moderate
Fatigue	-	-	Moderate
Difficulty focusing	-	-	Slight
Dizziness	Yes	-	Yes
Stomach awareness	Yes	Yes	Yes
Vertigo	-	-	Yes
Aware of breathing	-	-	Yes
Loss of appetite	-	-	Yes
Desire to move bowels	-	-	Yes
Confusion	-	-	Yes
Burping	-	-	Yes

Notes. 1. Only the symptoms that increased between the preflight and postflight 1 SSQs are listed.
2. The highest symptom severity level checked is provided.
3. Dash indicates the symptom was not present or did not increase.

Table 9. Simulator Sickness Symptomatology of the Pilots Who Were Unable to Continue the Sessions Using the LFOVD

Symptom	Pilot Group/ID No. IP/10	Mixed/1	Mixed/4
General discomfort	Moderate	Moderate	Slight
Fatigue	Slight	Slight	-
Eye strain	Slight	-	-
Salivation increased	Slight	-	-
Sweating	Moderate	Severe	Moderate
Nausea	Slight	Moderate	Slight
Difficulty concentrating	-	Slight	-
Dizziness	Yes	Yes	-
Stomach awareness	Yes	Yes	Yes
Vertigo	-	Yes	-
Loss of appetite	-	Yes	-
Desire to move bowels	-	Yes	-

Notes. Same as Table 8.

The symptoms that were retained were as follows:

1. General discomfort
2. Fatigue
3. Headache
4. Eyestrain
5. Difficulty focusing
6. Increased salivation
7. Sweating
8. Nausea
9. Difficulty concentrating
10. Fullness of the head
11. Blurred vision
12. Dizziness
13. Vertigo
14. Stomach awareness
15. Burping

The 12 checklist items that were excluded were not retained because in previous simulator sickness surveys they had failed to show a change from pre- to post-simulator exposure, they did not differ in terms of frequency and severity across simulators, or they gave misleading indications of simulator sickness.

A three-factor ANOVA was used to analyze the SSQ scores. Pilot group was the between-subjects factor and the within-subjects factors were visual system and time of administration. The analysis indicated that the main effect of time of SSQ administration attained statistical significance, $F(2,36) = 6.92$, $p < 0.01$. The mean total scores for the pre, post 1, and post 2 SSQs were respectively 106.07, 117.67, and 103.93. Post hoc comparisons of the SSQ means using the least significant difference (LSD) test revealed that the post 1 mean differed significantly ($p < 0.05$) from the pre and post 2 means. The differences between the pre and post 2 means were not significant. These means indicate that, on the average, there was an increase in simulator sickness symptomatology as a function of exposure to the simulation among the pilots who completed the sessions. In addition, the pilots recovered from the adverse effects of the simulation to approximately the pre-flight level within 30 min following the simulator sessions. None of the other main effects or interactions were significant. The absence of a significant visual system main effect and the absence of significant interactions involving the visual system factor indicate that the incidence and severity of simulator sickness did not vary as a function of exposure to the LFOVD and the DART. That is, even though simulator sickness was induced from exposure to the simulation, the level of simulator sickness was approximately the same for both visual simulation systems. Similarly, there were no significant differences in simulator sickness between groups.

The percentage of pilots in the two groups that reported an increase in simulator sickness symptomatology between the pre and post SSQ administrations are presented in Table 10 for both the DART and the LFOVD. This includes all the symptoms that increased in severity from none to slight or higher and from one level to a higher level, say, slight to moderate. Table 10 also presents the percentage of pilots that reported each of the symptoms.

Ataxia Tests

Postsession WOFEC and SOLEC tests were not administered to the pilots who were unable to complete the simulator sessions because the ethical treatment of the pilots would have been compromised had they been required to perform the tests when they were already not feeling well. Therefore, the postsession WOFEC and SOLEC tests were administered only to the pilots who were able to complete the sessions. A repeated measures ANOVA was conducted for both the WOFEC and SOLEC test data. Pilot group constituted the between-subjects factor; visual system and time of administration were the within-subjects factors.

The main effect of time of administration in the WOFEC data analysis was statistically significant, $F(2,36) = 4.04$, $p < 0.01$. The corresponding mean WOFEC scores for the pre, post 1, and post 2 tests were 6.19, 4.78, and 5.89. LSD tests of the differences between the WOFEC means indicated that the pre and post 1 means and the post 1 and post 2 means were significantly different ($p < 0.05$), whereas the pre and post 2 means did not differ significantly. The difference between the first and second means signifies that postural equilibrium, as measured by the WOFEC test, was adversely affected as a result of exposure to the simulation. The pilots recovered to essentially the preflight level within 30 min following the end of the simulator sessions, as witnessed by the relatively small difference between the first and third WOFEC means. The visual system and pilot group main effects were not significant, nor were any of the interactions, which signified that the extent of postural disequilibrium induced by the LFOVD and DART visual systems was very nearly the same and that there were no differences between groups. The analysis of the SOLEC data indicated that none of the main effects or interactions were significant.

Task Difficulty Ratings

The task difficulty ratings that were collected at the completion of the 20-min single-ship flights and 20-min formation flights were subjected to an ANOVA with repeated measures, which entailed pilot group as the between-subjects factor and both visual system and flight task as within-subjects factors. The ratings pertained only to the pilots who completed the flights; ratings were not obtained from the pilots who were unable to complete the sessions due to simulator sickness.

Table 10. Simulator Sickness Symptomatology of the Pilots Who Completed the Sessions

	DART		LFOVD	
	IPs (%)	Mixed (%)	IPs (%)	Mixed (%)
	(N = 15)	(N = 5)	(N = 15)	(N = 6)
Overall incidence	73.33	80.00	53.33	83.33
<u>Symptoms experienced</u>				
General discomfort	45.45	25.00	50.00	60.00
Fatigue	45.45	50.00	62.50	40.00
Headache	-	-	25.00	-
Eye strain	54.55		12.50	40.00
Difficulty focusing	9.09	50.00	75.00	20.00
Salivation increased	18.18	-	12.50	20.00
Salivation decreased	9.09	-	-	-
Sweating	36.36	50.00	37.50	40.00
Nausea	45.45	-	12.50	40.00
Difficulty concentrating	-	-	12.50	-
Blurred vision	-	-	-	20.00
Dizziness	18.18	25.00	25.00	-
Vertigo	9.09	-	-	-
Visual flashbacks	-	-	-	20.00
Awareness of breathing	-	25.00	-	-
Stomach awareness	18.18	-	12.50	40.00
Loss of appetite	-	-	12 50	-
Increased appetite	9.09	-	-	-
Burping	-	25.00	12.50	20.00

<u>Notes</u>. 1. Only the symptoms that increased between the preflight and postflight 1 SSQs are listed.
2. Dash indicates the symptom was not present or did not increase.

The analysis indicated that there was a significant interaction between pilot group and flight task, $F(1,18) = 6.12$, $p < 0.05$. The interaction is depicted in Figure 9, which shows that the IPs considered the single-ship flight more difficult than the formation flight, whereas the mixed pilots provided higher difficulty ratings for the formation flight. The main effects in the analysis were not statistically significant, and none of the other two-way interactions or the three-way interaction attained significance.

Motion History Questionnaire (MHQ)

The MHQs were scored using the scoring method described by Kennedy et al. (1988). Pearson product moment correlation

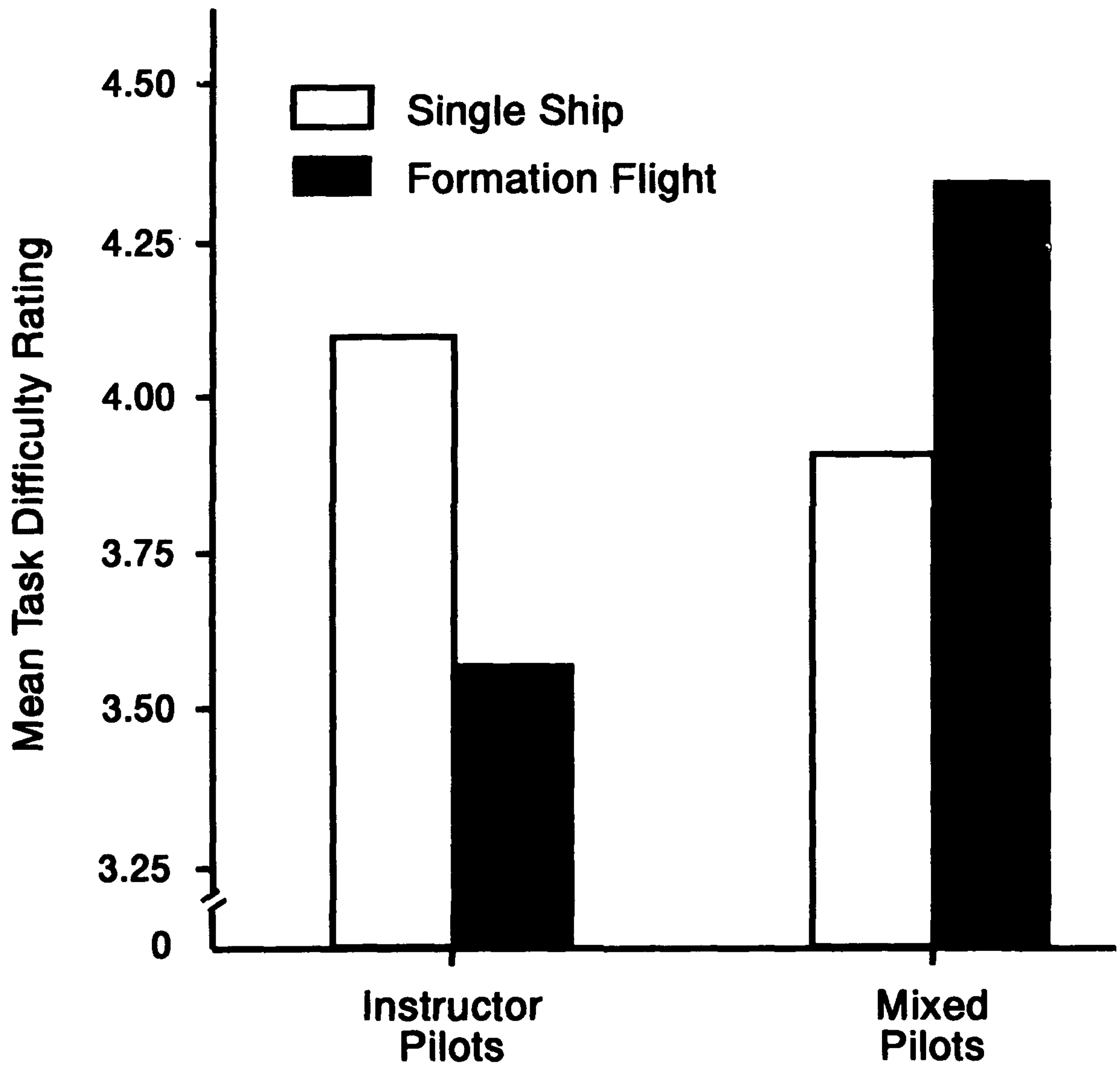

Figure 9
Pilot Task Difficulty Ratings for Single-Ship and Formation Flight Tasks

coefficients were subsequently obtained to assess the degree of association between the MHQ scores and the post 1 simulator sickness questionnaire (SSQ) scores. The SSQ scores and the MHQ scores that were paired in each of the computations are identified in Table 11, along with the corresponding Pearson coefficient (*r*). The table shows that none of the coefficients were statistically significant, which signifies that the post 1 SSQ scores were unrelated to the MHQ scores. It is imperative to point out that the pilot sample in the present investigation was quite small. Kennedy, et al. (1988) obtained low, albeit significant, correlations between MHQ scores and the incidence of simulator sickness in the evaluation of 10 U.S. Navy simulators when a very large pilot sample was used.

In-Flight Displays Evaluation

The pilots' responses to the display-related questions that were asked during the 20-min single-ship and formation flights are provided in Appendix A. One question was presented at 5 min, one at 10 min, and one at 15 min within each flight in the order they appear in the appendix. Each numerical value in the tables represents the percentage of pilots who gave the different responses. The responses of the pilots who terminated the sessions are included in the tables, up to the point where the sessions were ended. Thus, the number (N) of pilots from which feedback was obtained varies for each of the questions. For the most part, the pilot indicated that the simulated visual scene in the single-ship task and the lead aircraft in the formation flight task appeared realistic. Inspection of the pilots' responses suggests some changes that could be made to improve the fidelity of the simulation.

Steps could be taken, for example, to reduce or eliminate objects "popping" into the scene. In addition, the road in the canyon could be made darker to enhance its visibility, the trees could be scaled down, and the maneuvers of the lead aircraft could be made more predictable.

Display Evaluation Questionnaires

Pilot ratings for the DART and LFOVD display characteristics, along with the means and medians of the ratings, are provided in Appendix B. For the LFOVD, separate ratings were obtained for the AOI and background displays as well as for various overall display characteristics. The questionnaire was not administered to the pilots who prematurely exited the simulators due to severe discomfort because it was felt that they did not have sufficient experience with the visual systems to adequately rate the display features. It can be seen in Appendix B that, for the groups as a whole, none of the DART or LFOVD display characteristics were rated

Table 11. Correlations Between Post 1 SSQ Scores and MHQ Scores

No.	SSQ scores: Visual system	SSQ scores: Pilot group	MHQ scores: Pilot group	r	p
1	LFOVD	IPs	IPs	0.10	N.S.
2	LFOVD	Mixed	Mixed	0.28	N.S.
3	LFOVD	Combined IPs & mixed	Combined IPs & mixed	0.16	N.S.
4	DART	IPs	IPs	0.06	N.S.
5	DART	Mixed	Mixed	-0.10	N.S.
6	DART	Combined IPs & mixed	Combined IPs & mixed	-0.03	N.S.

Note. N.S. = Not significant; $p > 0.05$

below borderline and only two display features were considered borderline. These were the display resolution of the background display of the LFOVD and the vertical AOI excursion limits associated with the LFOVD.

Flight Performance Measures

The flight performance measures were subjected to both multivariate analysis of variance (MANOVA) and ANOVA statistical procedures. A three-factor model was used in these analyses, with visual system and flight time as the within-subjects factors and pilot group as the between-subjects factors. The flights were divided into five-min intervals to provide the flight time condition. Separate analyses were conducted for the single-ship flights and for the formation flights. The analyses encompassed only the pilots who were able to complete the simulator sessions. The simulator familiarization and the formation practice data were omitted. The results of these analyses are presented below.

Single-Ship Flight. The MANOVA for the single-ship task included all of the flight performance measures shown in Table 4 except the slant range mean and standard deviation. The results indicated that two of the main effects were statistically significant: (a) visual system, Wilks' $\lambda = 0.0019$, approximate $F(16,3) = 99.31$, $p < 0.01$, and (b) flight time, Wilks' $\lambda = 0.0237$, approximate $F(48,117) = 6.12$, $p < 0.01$. These findings signify that flight performance, as a whole, differed significantly between the

DART and LFOVD visual systems and as a function of flight duration. The main effects of pilot group and the interactions were not significant.

The significant main effects and interactions observed in the ANOVAs are provided in Table 12, and the corresponding means are presented in Appendix C. Table 12 shows that the main effect of flight time was statistically significant in all the ANOVAs, the main effect of visual system was significant for various flight performance measures, and none of the pilot group main effects were significant. The interactions involving pilot group were significant, however, for several performance measures. There were no significant two-way interactions between visual system and flight time. This suggests that the performance differences over time were approximately the same for both visual systems when the data were pooled for the two pilot groups.

Formation Flight. The MANOVA conducted for the formation flight task encompassed all of the performance measures listed in Table 4, except the slant range mean and standard deviation. The latter two scores were not included in this analysis due to the limited degrees of freedom in the research design. The slant range measures are addressed later in this section.

The results of the MANOVA revealed that the main effect of flight time was significant, Wilks' $\lambda = 0.0007$, approximate $F(48,117) = 25.53$, $p < 0.01$, as was the visual system by flight time interaction, Wilks' $\lambda = 0.091$, approximate $F(48,105) = 2.71$, $p < 0.01$. From these results, it is evident that overall flight performance in the formation flight task varied with respect to flight time and that the changes in performance over time were dependent upon the particular visual system used.

The significant main effects and interactions obtained in the ANOVAs are identified in Table 13, and the corresponding means are provided in Appendix D. The univariate analyses for slant range are presented in Table 13 to facilitate comparison of the performance measurement ensemble.

Disregarding slant range, Table 13 shows that the flight time main effect was significant for all of the performance measures except the mean absolute and the standard deviation pitch angle. In addition, the main effect of visual system was significant for a number of performance measures, but the main effect of pilot group was not significant in any of the univariate analyses associated with the formation flights. Three of the two-way visual system by flight time interactions were significant, which indicated that in some cases the variations in flight performance over time were different between the two visual systems. None of the interactions involving pilot group were significant, suggesting that the two pilot groups performed the formation flights about the same way.

Table 12. Significant Main Effects and Interactions in the Univariate Analyses of the Single-Ship Flight Performance Measures

Flight Performance Measure		VS	FT	PG	VSxFT	VSxPG	FTxPG	VSxFTxPG
G-force	Mean	*	**					
	Std. Dev.	**	**					
Altitude	Mean	**	**			*	*	
	Std. Dev.		*			**		
Airspeed	Mean		**					*
	Std. Dev.	*	**				*	*
Pitch angle	Mean Abs.		**			*		
	Std. Dev.		**			**		
Bank angle	Mean Abs.		**					
	Std. Dev.	*	**					
Pitch rate	Mean Abs.	**	**					
	Std. Dev.	**	**					
Roll rate	Mean Abs.	**	**				*	
	Std. Dev.	**	**					
Yaw rate	Mean Abs.		**					
	Std. Dev.	**	**					

Notes. 1. VS = visual system.
2. FT = flight time.
3. PG = pilot group.
4. * = $p < 0.05$.
5. ** = $p < 0.01$.

Table 13. Significant Main Effects and Interactions in the Univariate Analyses of the Formation Flight Performance Measures

Flight Performance Measure		VS	FT	PG	VSxFT	VSxPG	FTxPG	VSxFTxPG
G-force	Mean		**					
	Std. Dev.	**	**		**			
Altitude	Mean		**					
	Std. Dev.		**					
Airspeed	Mean	**	**					
	Std. Dev.		**					
Pitch angle	Mean Abs.							
	Std. Dev.							
Bank angle	Mean Abs.		**					
	Std. Dev.		**					
Pitch rate	Mean Abs.	**	**		**			
	Std. Dev.	**	**		**			
Roll rate	Mean Abs.	**	**					
	Std. Dev.	**	**					
Yaw rate	Mean Abs.	**	**					
	Std. Dev.	**	**		*			
Slant range	Mean					*		
	Std. Dev.				*	*		

<u>Notes</u>. 1. VS = visual system.
2. FT = flight time.
3. PG = pilot group.
4. * = $p < 0.05$.
5. ** = $p < 0.01$.

An additional MANOVA was conducted that encompassed the two slant range scores for the formation flights, i.e., mean slant range and standard deviation slant range. In this analysis, the main effect of flight time was significant, Wilks' $\lambda = 0.7568$, approximate $F(6,106) = 2.64$, $p < 0.05$. None of the remaining main effects nor any interactions were significant.

The results of the ANOVAs are presented in Table 13, and the means of the significant interactions are in Appendix D. The significant two-way visual system by pilot group interactions that were observed for both the mean and standard deviation slant range indicate that the effects of the two visual systems differed between the two pilot groups. Additionally, the significant visual system by flight time interaction for the standard deviation slant range shows that slant range deviations varied as a function of time at the task.

Relationship Between Symptomatology and Flight Performance

Pearson coefficients were obtained to determine the extent of the relationship between the post 1 SSQ scores and the flight performance scores for both the DART and the LFOVD. These coefficients are presented in Table 14. The performance scores consisted of the average of the mean scores for the four intervals within the last flight the pilots performed in a session. That is, for the pilots who performed the 20-min formation flight followed by the 20-min single-ship flight, the means of the four 5-min intervals in the 20-min single-ship flight were averaged. Conversely, for the pilots who performed the single-ship flight first and then the formation flight, the means of the four 5-min intervals in the formation flight were averaged. This approach was adopted to permit a comparison of the correlation coefficients for the two types of tasks. The scores for both the IPs and the mixed group of pilots were included in the computation of each coefficient. For those pilots who had to terminate the sessions due to severe simulator sickness, the scores from the flight they were performing when the simulation was suspended were used.

Table 14 shows that 8 of the 18 flight performance measures were significantly related to the post 1 SSQ scores. Moreover, it can be seen that there were both positive and negative correlations. This indicates that in some instances, an increase in the flight performance scores was associated with an increase in the SSQ scores and in other instances, decreasing performance scores were associated with increasing SSQ scores. The manner in which the significant coefficients are distributed in Table 14 indicates that the relationships between the post 1 SSQ scores and the flight performance measures were dependent upon both the visual system and performance task. Only one significant correlation was obtained in relation to the DART, which occurred with the formation task. Seven significant correlations were observed in conjunction with the LFOVD, two with the single-ship task and five with the formation task.

Table 14. Relationships Between Post 1 SSQ Scores and Flight Performance Scores

Flight performance measure	Visual System: DART, Last flight in session: Single ship	DART, Last flight in session: Formation	LFOVD, Last flight in session: Single ship	LFOVD, Last flight in session: Formation
Mean G-force	-0.20	-0.48	-0.13	-0.58*
S.D. G-force	-0.12	0.55	0.12	-0.41
Mean altitude	0.31	-0.04	0.40	0.64*
S.D. altitude	0.001	0.01	0.47	0.48
Mean airspeed	-0.001	0.14	-0.10	-0.67*
S.D. airspeed	0.27	0.22	0.12	0.14
Mean absolute pitch angle	0.12	-0.17	0.58*	0.20
S.D. pitch angle	-0.06	-0.22	0.65*	0.10
Mean absolute bank angle	-0.15	-0.49	0.07	-0.22
S.D. bank angle	-0.20	-0.45	0.06	-0.33
Mean absolute pitch rate	-0.06	0.55	0.15	-0.49
S.D. pitch rate	-0.03	0.63*	0.16	-0.36
Mean absolute roll rate	-0.24	-0.22	0.07	-0.66*
S.D. roll rate	-0.21	-0.19	0.01	-0.59*
Mean absolute yaw rate	0.06	-0.25	0.22	0.05
S.D. yaw rate	0.12	-0.16	0.17	-0.14
Mean slant range	-	-0.17	-	0.52
S.D. slant range	-	-0.24	-	0.46

Notes. 1. Values represent Pearson coefficients (r).
2. * = $p < 0.05$.

Relationship Between Discomfort Ratings and Flight Performance

Tables 15 and 16 show the degree of association between the in-flight pilot discomfort ratings and the corresponding flight performance scores obtained with the DART and LFOVD, respectively. The discomfort ratings for both the IPs and the mixed group of pilots were pooled. The coefficients presented at the 5- and 10-min time periods correspond to the simulator familiarization flights and the formation practice flights. Since the 20-min single-ship and formation flights were counterbalanced, each of the time intervals from 15 min to 50 min encompasses both flights. The ratings of the pilots who had to terminate the simulation were included up to the interval when the sessions were terminated.

Fifteen of the 18 flight performance measures were significantly related to the pilot discomfort ratings, and both positive and negative correlations were obtained. Examination of Tables 15 and 16 suggests that the degree of association between the scores was a function of the visual system and flight time. There were 8 significant correlations in conjunction with the DART and 36 with the LFOVD. For the DART, the significant correlations occurred with the formation practice flight (represented by the 10-min period in the tables) and in the last 5 min of the initial 20-min flight (represented by the 30-min period in the tables). Table 16 shows that the significant correlations associated with the LFOVD occurred throughout the various time periods but primarily in the last 5 min of the initial 20-min flight.

Flight Resets

The frequency of resets in the single-ship flights due to terrain crashes and cloud penetrations is shown in Table 17, and the resets due to terrain crashes, cloud penetrations, and manual re-forms (i.e., researcher-initiated resets) are presented in Table 18 for the formation flights. In both flights there were more terrain crashes than cloud penetrations, and only four manual re-forms were required in the formation flights. The terrain crashes were more prevalent because the pilots were required to fly low and perform vigorous maneuvers. Both the IPs and mixed group of pilots contacted the terrain and penetrated the clouds, and the resets occurred in the DART as well as the LFOVD. The number of terrain crashes is substantially higher for the IPs than the mixed group, but this was mainly due to a couple of IPs who occasionally made contact with the trees, which were counted as terrain crashes.

Table 15. Relationships Between In-Flight Discomfort Ratings and Flight Performance Scores Associated with the DART

Flight performance measure	Flight time (minutes) 5	10	15	20	25	30	35	40	45	50
Mean G-force	-0.10	-0.13	-0.19	-0.08	-0.02	-0.12	-0.06	0.28	0.04	-0.19
S.D. G-force	0.04	0.40	0.33	0.03	0.22	0.10	0.05	0.12	0.13	-0.26
Mean altitude	0.10	0.09	0.21	0.05	-0.12	-0.09	-0.08	0.002	0.22	0.03
S.D. altitude	0.22	0.07	0.13	0.08	-0.09	-0.02	-0.08	-0.04	0.09	-0.17
Mean airspeed	-0.09	0.04	0.03	0.05	0.11	-0.31	-0.26	-0.08	0.08	-0.22
S.D. airspeed	0.21	0.55*	0.09	0.009	0.07	0.02	-0.07	0.23	0.35	0.24
Mean absolute pitch angle	0.32	-0.09	0.08	0.07	-0.07	0.05	0.05	0.15	0.13	-0.08
S.D. pitch angle	0.34	-0.15	0.05	0.08	-0.06	0.10	0.04	0.06	0.10	-0.13
Mean absolute bank angle	0.09	0.27	-0.07	-0.07	-0.09	-0.02	0.04	0.18	0.15	-0.08
S.D. bank angle	0.003	0.07	-0.04	0.02	-0.11	0.08	0.002	0.15	0.11	-0.05
Mean absolute pitch rate	0.10	0.46*	0.34	0.04	0.19	0.38	0.15	0.29	0.12	-0.11
S.D. pitch rate	0.09	0.49*	0.40	0.08	0.30	0.49*	0.25	0.26	0.16	-0.06
Mean absolute roll rate	0.06	0.19	-0.07	0.07	0.19	0.51*	0.10	0.30	0.09	-0.08
S.D. roll rate	-0.003	0.21	-0.12	0.06	0.22	0.57**	0.09	0.34	0.19	-0.04
Mean absolute yaw rate	0.12	0.22	-0.08	0.02	0.24	0.54*	0.40	0.45	0.41	-0.03
S.D. yaw rate	0.05	0.12	-0.18	0.09	0.32	0.60**	0.36	0.45	0.40	-0.03
Mean slant range	-	0.11	0.31	-0.28	-0.13	-0.09	0.31	-0.17	0.27	-0.17
S.D. slant range	-	0.18	0.20	-0.17	0.07	-0.21	0.39	0.18	0.23	-0.36

Notes. 1. Values represent Pearson coefficients (r).
2. Single-ship and formation flights are combined from 15 to 50 minutes.
3. * = $p < 0.05$.
4. ** = $p < 0.01$.

Table 16. Relationships Between In-Flight Discomfort Ratings and Flight Performance Scores Associated with the LFOVD

Flight performance measure	Flight time (minutes) 5	10	15	20	25	30	35	40	45	50
Mean G-force	-0.12	0.05	-0.07	0.19	0.02	-0.59**	-0.05	-0.17	-0.40	0.36
S.D. G-force	0.01	0.37	0.02	0.37	0.30	-0.10	0.14	-0.16	-0.10	0.31
Mean altitude	0.40*	0.42*	0.45*	0.54**	0.69**	0.63**	-0.11	-0.27	-0.15	-0.15
S.D. altitude	0.38	0.55**	0.35	0.47*	0.54**	0.62**	-0.14	-0.30	-0.22	-0.12
Mean airspeed	0.23	-0.09	0.12	0.47*	0.22	-0.05	-0.28	-0.48*	-0.20	-0.19
S.D. airspeed	0.10	0.17	0.32	0.08	0.42*	0.23	0.10	0.29	-0.07	0.13
Mean absolute pitch angle	0.29	0.51*	0.31	0.55**	0.45*	0.46*	-0.02	-0.06	-0.24	0.03
S.D. pitch angle	0.28	0.48*	0.28	0.53**	0.37	0.40	-0.06	0.07	-0.26	-0.004
Mean absolute bank angle	0.18	0.25	-0.03	0.41	0.35	-0.21	0.21	-0.24	-0.21	0.25
S.D. bank angle	0.14	0.33	-0.04	0.41	0.34	-0.36	0.17	-0.30	-0.29	0.11
Mean absolute pitch rate	-0.02	0.22	-0.16	0.05	0.07	-0.54**	0.37	0.31	-0.14	0.58**
S.D. pitch rate	-0.02	0.44*	-0.07	0.02	0.11	-0.33	0.42	0.39	0.17	0.62**
Mean absolute roll rate	-0.02	0.25	-0.25	0.01	-0.13	-0.57**	0.27	-0.12	-0.20	0.33
S.D. roll rate	-0.01	0.33	-0.19	-0.02	-0.04	-0.53**	0.17	-0.03	-0.12	0.20
Mean absolute yaw rate	0.05	0.03	-0.21	-0.24	-0.01	-0.36	0.45**	0.45*	0.34	0.61**
S.D. yaw rate	-0.06	0.14	-0.25	-0.31	-0.09	-0.49*	0.40	0.45*	0.09	0.56**
Mean slant range	-	0.09	0.48	-0.04	0.13	0.80**	-0.24	0.12	0.32	-0.39
S.D. slant range	-	-0.01	0.60*	0.01	0.78**	0.74**	-0.38	0.10	0.39	-0.26

Notes. 1. Values represent Pearson coefficients (r).
2. Single-ship and formation flights are combined from 15 to 50 minutes.
3. * = $p < 0.05$.
4. ** = $p < 0.01$.

Table 17. Frequency of Terrain Crashes and Cloud Penetrations for Single-Ship Flights

Visual system	Pilot group	Terrain crashes Flight time (min) 5	10	15	20	Cloud penetrations Flight time (min) 5	10	15	20
DART	Ips	1	4	4	6	0	0	0	0
	Mixed	0	0	0	2	1	0	0	0
LFOVD	IPs	5	7	4	5	0	1	0	0
	Mixed	1	1	0	3	0	0	0	0

Note. Includes only the pilots who were able to complete the flights.

Table 18. Frequency of Terrain Crashes, Cloud Penetrations, and Manual Re-Forms for Formation Flights

Visual system	Pilot group	Terrain crashes Flight time (min) 5	10	15	20	Cloud penetrations Flight time (min) 5	10	15	20	Manual re-forms Flight time (min) 5	10	15	20
DART	IPs	4	1	5	4	1	1	0	4	1	0	0	1
	Mixed	1	3	0	3	0	1	2	0	0	0	0	1
LFOVD	IPs	1	3	3	9	1	2	3	0	0	0	1	0
	Mixed	1	0	0	2	0	0	0	0	0	0	0	0

Note. Includes only the pilots who were able to complete the flights.

DISCUSSION

Simulator Sickness

The present investigation demonstrated that both the DART and LFOVD visual simulation systems can induce simulator sickness symptomatology in low-level, high-speed flight situations involving rigorous aircraft maneuvering. Several pilots were prematurely forced to terminate the simulator sessions with both visual systems due to severe discomfort. Even among the pilots who were able to complete the sessions, there was a significant increase in self-reports of pilot discomfort and simulator sickness symptoms as a function of exposure to the visual systems. In addition, the visual systems had a significant adverse effect on post-flight postural equilibrium, which dissipated after 30 min.

Both the DART and the LFOVD produced approximately the same level of simulator sickness among the pilots in terms of the discomfort ratings, simulator sickness symptoms, and postural disequilibrium. The number of pilots that were unable to complete the sessions was also nearly the same. Only one pilot that was unable to complete the session using the DART finished the entire session with the LFOVD, but at the end of the session when the LFOVD was used, the pilot exhibited severe sweating and pallor, which are prominent overt signs of simulator sickness. This pilot indicated at the conclusion of the session with the LFOVD that he was determined to complete the simulator flights.

The severest cases of simulator sickness, where the pilots suspended the simulation before the end of the sessions, were more pronounced among the "mixed" group of pilots who were older and had more flight experience, on the average, than the IPs. In the mixed group, three of the eight pilots (37.5%) became too ill to finish the flights with the LFOVD and/or the DART. Only one of the 16 IPs (6.25%), on the other hand, failed to complete the sessions. Coincidentally, this IP was the only female participant, and she indicated that she was very susceptible and had succumbed to motion sickness during other flight simulation. The analyses showed that there were no differences in simulator sickness between the two groups of pilots that were able to complete the simulator sessions.

The relationship between flight experience and the incidence and severity of simulator sickness has been evaluated in a number of previous investigations. In some, simulator sickness occurrences were found to increase as pilot flight experience increased (Braithwaite & Braithwaite, 1990; Crowley, 1987; Kennedy, Merkle, & Lilienthal, 1985; McGuinness, Bouwman, & Forbes, 1981). An explanation for this relationship is that the sensory cues presented to the pilot during the simulation are inconsistent with past neural stores acquired through experience in actual flight. This inconsistency constitutes the basic foundation of the "sensory conflict" theory or "neural mismatch" theory, which is the most common theoretical model used to explain simulator sickness (Benson, 1988; McCauley, 1984; Reason, 1978). It should be pointed out, however, that in some instances no significant relationship between flight experience and simulator sickness was observed (Chappelow, 1988; Magee, Kantor, & Sweeney, 1988; Ungs, 1988).

In this investigation, the self reports of the incidence and severity of simulator sickness symptomatology obtained through the application of the simulator sickness questionnaire (SSQ) were converted to a single, total score using a procedure based on the SSQA scoring method described by Lane and Kennedy (1988). Comparison of the total SSQA scores in the present research with the scores derived by Lane and Kennedy for nine U.S. Navy flight simulators indicates that the reported symptomatology associated with both the DART and LFOVD visual simulation systems was relatively high. Table 19 shows the mean total SSQA scores of the individual and combined pilot groups in relation to the DART and LFOVD. This table presents only the post 1 (immediately following the simulator sessions) SSQA scores and it includes the pilots who voluntarily terminated the simulator sessions. In contrast with the scores in Table 19, the highest mean total SSQA score obtained by Lane and Kennedy was 118.8. The magnitude of the scores associated with the visual systems compared in this evaluation suggests that the DART and LFOVD are highly provocative and that further research is warranted to determine the sources of the problems and the means to correct them. It is necessary to point out, however, that the Lane and Kennedy data were based on considerably more pilots and simulator flights than in the present investigation, and their sample would be more representative of the true pilot population. Thus, before extensive steps are undertaken to reduce simulator sickness induced by the DART and LFOVD, follow-on research should be conducted employing a larger sample size that is representative of the anticipated user population.

Although a relatively large number of the pilots who were able to complete the sessions reported an increase in simulator sickness symptomatology between the pre and post 1 SSQ administrations (see Table 10), self reports of discomfort during the sessions increased from a mean of 1.24 to only 1.73 on the 7-point discomfort scale. This suggests that the pilots were consciously aware of existing symptoms, but the symptoms produced little discomfort, on the average, during the sessions. McCauley, et al. (1990) also observed a similar, small increase in discomfort over a simulated 40-min helicopter-following task for the experimental condition in which the simulator motion was deactivated. The mean discomfort ratings ranged from a minimum of about 1.45 to a maximum of approximately 2.6 on a 7-point scale. These means encompassed both the ratings of the pilots who had to terminate the simulation and the pilots who completed the flight sessions.

The in-flight display-related questions and the display evaluation questionnaires gave no indication that there were any

Table 19. Mean Total SSQA Scores for the Individual and Combined Groups in Relation to the DART and LFOVD

Pilot group	Mean total SSQA score	
	DART	LFOVD
IPs	121.97	121.97
Mixed	140.23	134.19
Combined IPs & mixed	128.05	125.69

serious display deficiencies that may have provoked the simulator sickness incidents. Although the questionnaire approach is not absolutely conclusive, it would appear that the simulator sickness symptoms were induced by factors other than display deficiencies. Some of the more likely causes of the simulator sickness occurrences were: the susceptibility of the pilots, the wide-angle visual scenes, the richness of the scene content, the closeness to the ground in the flights, the intense maneuvering, and the duration of the flights. Moreover, it is proposed that these factors interacted to produce the characteristic effects on the pilots that were observed, such that one factor without or with lesser values of the other factors may not have provoked the same amount of simulator sickness. For example, susceptible pilots may not experience any symptoms if they perform straight-and-level flights at a simulated altitude of 25,000 feet over flat terrain.

Simulator Sickness Countermeasures

Because both the DART and LFOVD visual simulation systems can induce simulator sickness, countermeasures should be considered in the use of these systems to avoid or minimize the occurrence and severity of the malaise and the potentially adverse effects of the simulation on the health and safety of the users and user performance. A variety of guidelines for alleviating simulator-induced sickness have been proposed (Casali, 1986; Kennedy et al., 1987; McCauley, 1984). The guidelines that are applicable to the DART and LFOVD visual systems are presented in Appendix E. Some of these were adopted in the design of the present investigation in an effort to minimize the number of variables that might have contributed to simulator sickness. For example, the visual scene was blanked with a cloud-like mask when the pilots ingressed and egressed the cockpit and when the simulator was frozen.

The pilot discomfort ratings that were obtained during the flights indicated that the pilots recovered slightly from the ill effects of the simulation during the flights when the tasks were changed. This suggests that the tasks should be varied as much as possible to prevent the onset of simulator sickness symptomatology and the tasks should be of relatively short duration. An alternative countermeasure might be to intersperse mild maneuvering tasks that are flown at high altitudes among the more difficult tasks.

The significant decline in postural equilibrium that was observed, as measured by the Walk on Floor Eyes Closed (WOFEC) test, suggests that some personnel may encounter balance difficulties following the use of the DART and LFOVD visual simulation systems. For this reason, they should not be allowed to leave the area immediately after using the simulators. The research and simulator staff should walk with them around the facility to observe whether any balance problems are evident and to allow them to recover. The post 2 WOFEC test indicated that the pilots as a whole had recovered from the adverse effects of the simulation 30 min after the sessions ended.

Effects of the Treatment Conditions on Flight Performance

The analyses revealed that the flight performance of the pilots who were able to complete the sessions differed significantly between the LFOVD and the DART in both the single-ship and formation flight tasks. Inspection of the means associated with the significant visual system main effects in Appendix C and D indicates that the pilots exhibited a tendency to overcontrol the DART F-16C simulator compared to the LFOVD F-16A simulator. It can be seen, for example, that the pitch rates and roll rates (both absolute means and standard deviations) were consistently higher for the DART F-16C.

The differences in pilot performance that were observed between the DART F-16C and LFOVD F-16A simulators cannot be attributed to differences in simulator sickness. This is because the analyses showed that the incidence and severity of simulator sickness, as measured with the SSQs, discomfort ratings, and the ataxia tests, were approximately the same (i.e., not statistically different) for both simulators. It is surmised that the differences in pilot performance between the simulators were largely due to the more sensitive side stick and the greater transport delay associated with the DART F-16C simulator.

Flight performance also varied significantly as a function of time at the task. For the single-ship task, the means associated with the significant main effects of flight time remained relatively constant or declined slightly up to the 15-min point of the task and then increased markedly in the last 5 min. A reasonable hypothesis for these trends is that the level of task difficulty was about the same and the pilots became more proficient in flying the simulator up to the 15-min point. During the last 5 min, however, the canyon in which they were flying became narrower and more turns were required, which elevated the difficulty of the task. In addition to the higher workload, the pilots may have experienced debilitating fatigue and simulator sickness symptomatology, which would have adversely influenced their flight performance.

In the formation task, flight performance was largely dictated by the maneuvers of the prerecorded lead aircraft. Thus, because the lead aircraft's maneuvers were intentionally varied throughout the 20-min formation flight, the means of the flight performance measures associated with the significant main effects of flight time also varied. Pilot fatigue, control proficiency, and simulator sickness may have also contributed to the performance differences that were observed over time.

There was no significant interaction between visual system and flight time for any of the performance measures in the single-ship task, which indicated that the variations in flight performance over time were similar for both visual systems. The visual system and pilot group conditions interacted, however, with respect to the mean and standard deviation altitude and the mean absolute and standard deviation pitch angle performance measures for the single-ship task. In Table 3 of Appendix C, it is evident that both the IPs and the mixed group of pilots flew lower to the ground on the average with the LFOVD than with the DART, and the IPs flew closer with both visual systems when compared to the mixed group. Table 4 in Appendix C shows, however, that the standard deviation altitude measure was greater for the IPs in relation to the LFOVD and greater for the mixed group in conjunction with the DART. This suggests that there may be a "ceiling/floor" effect, where the closer the pilots fly to the limits of the allowable vertical excursion, the greater the variation in altitude to avoid contacting the ground or exceeding the upper limit. The excursion limits were defined by the terrain at one extreme and the flashing altitude indicator lights in the cockpit at the other extreme when the pilot exceeded 500 feet. In the single-ship flights, the IPs flew closer to the terrain and the mixed pilots flew closer to the upper excursion limit. The mean absolute and standard deviation pitch angles exhibit the same trend (see Tables 7 and 8 in Appendix C) as described above for standard deviation altitude, which is consistent with the interrelationship between changes in pitch angle and altitude variation.

Several two-way flight time by pilot group interactions were also significant for the single-ship task, which basically indicates that the flight performance within and between the two groups of pilots varied over the duration of the flights. Various factors could have accounted for these differences, such as differences in simulator proficiency, fatigue, simulator sickness, and task difficulty. Significant three-way interactions were also observed in the single-ship flights in relation to airspeed. This essentially means that airspeed varied within and between the two pilot groups over the period of the task, and that the variations in airspeed differed between the two visual systems.

In the analyses of the formation flight performance measures, several two-way visual system by flight time interactions were significant. As mentioned previously, the differences in performance over time were due to the variations in the maneuvers performed over the route by the lead aircraft. In each of the interactions except for slant range, the magnitude of the performance means associated with the DART were greater than the means corresponding to the LFOVD, which can be seen in Tables 2, 9, 10, and 14 of Appendix D. The larger means are indicative of more pronounced stick inputs.

The two-way interaction between visual system and flight time that was observed for the standard deviation slant range scores indicate that the pilots initially had much more difficulty maintaining position with the simulated lead aircraft in the DART F-16C simulator than in the LFOVD F-16A simulator. The differences between the slant range scores diminished, however, as the flight progressed, and the slant range scores obtained with the DART F-16C simulator were even slightly smaller than the scores associated with the LFOVD F-16A simulator in the last half of the formation flights, which is evidenced by the mean scores presented in Table 16 of Appendix D. One explanation for these findings is that due to the greater stick sensitivity and transport delay, the pilots mainly focused on controlling the aircraft in the DART F-16C simulator at the expense of slant range separation. Consequently, the pilots were unable to stabilize slant range during the initial period of the formation flights as well as they could with the LFOVD F-16A simulator. As the pilots became more proficient at flying the DART F-16C simulator as the mission progressed, they were able to stabilize slant range at about the same extent as in the LFOVD F-16A simulator.

Slant range separation also varied as a function of the differences between the two pilot groups. In the analysis, a significant interaction between visual system and pilot group was obtained for both the mean and the standard deviation slant range performance measures. The averages of the mean slant range scores provided in Table 15 in Appendix D show that the IPs maintained closer formation than the mixed pilots with both the DART and LFOVD and that the mixed pilots flew closer formations with the LFOVD than the DART. Table 16 in Appendix D indicates that the average of the standard deviation slant range scores was substantially higher for the mixed pilots when using the DART. This suggests that more sensitive control stick and transport delay associated with the DART F-16C simulator adversely affected the mixed pilots more than the IPs.

The analyses showed that flight performance was related to the level of simulator sickness the pilots experienced during the simulator sessions. Significant positive and negative correlations were observed between the post 1 SSQ scores and 8 of the 18 flight performance measures. In addition, significant positive and negative correlations were obtained between the discomfort ratings and 15 of the performance measures. These correlations may be interpreted in several ways. One is that they could signify that simulator sickness influenced pilot flight performance during the simulator sessions. Conversely, they may also indicate that simulator sickness was governed by how the pilots maneuvered the aircraft in the flights. In addition, both events may have occurred concurrently across the various pilots. That is, for some pilots, simulator sickness may have produced a change in performance while for other pilots, the manner in which they performed the simulated flights may have induced the simulator sickness.

IMPLICATIONS

Simulator sickness symptomatology was experienced by a majority of the pilots in conjunction with the use of both the DART and LFOVD visual simulation systems in the present investigation. Some pilots experienced severe distress and were forced to terminate the simulator sessions, while others were able to complete the flights. Had the simulator sickness symptomatology among this latter group of pilots not been specifically addressed, their symptoms may have escaped entirely unnoticed. The implication of this is that many pilots involved in flight simulation may experience simulator sickness without anyone's knowledge, since tests of simulator sickness are typically not administered.

REFERENCES

Baltzley, D.R., Kennedy, R.S., Berbaum, K.S., Lilienthal, M.G., & Gower, D.W. (1989). The time course of postflight simulator sickness symptoms. Aviation, Space, and Environmental Medicine, 60, 1043-1048.

Benson, A.J. (1988). Aetiological factors in simulator sickness (AGARD No. 433, AD-A202 492). In: Motion cues in flight simulation and simulator induced sickness, AGARD Conference Proceedings. Neuilly Sur Seine, France: North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development.

Braithwaite, M.G., & Braithwaite, B.D. (1990). Simulator sickness in an Army simulator. Journal of the Society of Occupational Medicine, 40, 105-110.

Casali, J.G. (1986). Vehicular simulation-induced sickness, Volume I: An overview (NTSC TR 86-010, AD-A173 904). Orlando, FL: Naval Training Systems Center.

Casali, J.G., & Frank, L.H. (1988). Manifestation of visual/ vestibular disruption in simulators: Severity and empirical measurement of symptomatology (AGARD No. 433, AD- A202 492). In: Motion cues in flight simulation and simulator induced sickness, AGARD Conference Proceedings. Neuilly Sur Seine, France: North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development.

Chappelow, J.W. (1988). Simulator sickness in the Royal Air Force: A survey (AGARD No. 433, AD-A202 492). In: Motion cues in flight simulation and simulator induced sickness, AGARD Conference Proceedings. Neuilly Sur Seine, France: North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development.

Crowley, J.S. (1987). Simulator sickness: A problem for Army aviation. Aviation, Space, and Environmental Medicine, 58, 355-357.

Fregly, A.R., Graybiel, A., & Smith, M.J. (1972). Walk on floor eyes closed (WOFEC): A new addition to an ataxia test battery. Aerospace Medicine, 43(4), 395-399.

Hamilton, K.M., Kantor, L., & Magee, L.E. (1989). Limitations of postural equilibrium tests for examining simulator sickness. Aviation, Space, and Environmental Medicine, 60, 246-251.

Howard, C.M. (1989). Display characteristics of example light-valve projectors (AFHRL-TR-88-44, AD A209 580). Williams AFB, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Kennedy, R.S., Berbaum, K.S., Allgood, G.O., Lane, N.E., Lilienthal, M.G., & Baltzley, D.R. (1988). Etiological significance of equipment features and pilot history in simulator sickness (AGARD No. 433, AD-A202 492). In: Motion cues in flight simulation and simulator induced sickness, AGARD Conference Proceedings. Neuilly Sur Seine, France: North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development.

Kennedy, R.S., Berbaum, K.S., Lilienthal, M.G., Dunlap, W.P., Mulligan, B.E., & Funaro, J.F. (1987). Guidelines for alleviation of simulator sickness symptomatology (NAVTRASYSCEN TR-87-007). Orlando, FL: Naval Training Systems Center.

Kennedy, R.S., Lilienthal, M.G., Berbaum, K.S., Baltzley, D.R., & McCauley, M.E. (1989). Simulator sickness in U.S. Navy flight simulators. Aviation, Space, and Environmental Medicine, 60, 10-16.

Kennedy, R.S., Merkle, Jr., P.J., & Lilienthal, M.G. (1985). A comparison of postural equilibrium effects following exposure to different ground-based flight trainers. In: Proceedings of the 23rd Annual Symposium of the SAFE Association. Van Nuys, CA: SAFE Association.

Lane, N.E., & Kennedy, R.S. (1988). A new method for quantifying simulator sickness: Development and application of the simulator sickness questionnaire (SSQ) (EOTR 88-7). Orlando, FL: Essex Corp.

Leinenwever, R.W., & Moran, S.I. (1992). Transport delay measurements: Methodology and analysis for the F-16C Combat Engagement Trainer, the Display for Advanced Research and Training, and the F-16A Limited Field of View (AL-TP-1992-0009). Williams AFB, AZ: Aircrew Training Research Division, Human Resources Directorate.

Magee, L.E., Kantor, L., & Sweeney, D.M.C. (1988). Simulator induced sickness among Hercules aircrew (AGARD No. 433, AD-A202 492). In: Motion cues in flight simulation and simulator induced sickness, AGARD Conference Proceedings. Neuilly Sur Seine, France: North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development.

McCauley, M.E. (Ed.) (1984). Research issues in simulator sickness: Proceedings of a workshop. Washington, DC: National Academy Press.

McCauley, M., Hettinger, L., Sharkey, T., & Sinacori, J. (1990). The effects of simulator visual-motion asynchrony on simulator induced sickness (AIAA-90-3172). Washington, DC: American Institute of Aeronautics and Astronautics.

McGuinness, J., Bouwman, J.H., & Forbes, J.M. (1981). Simulator sickness occurrences in the 2E6 Air Combat Maneuvering Simulator (ACMS) (NAVTRAEQUIPCEN 80-C-0135-4500-1). Orlando, FL: Naval Training Equipment Center.

Meister, D. (1986). Advances in Human Factors/Ergonomics, 5: Human Factors Testing and Evaluation. New York: Elsevier Science Publishing Company, Inc.

Parfitt, A.G., & Chappelow, J. (1986). A survey of simulation sickness amongst Royal Air Force pilots. Report on interim results. In: Advances in flight simulation - Visual and motion systems, International Conference Proceedings. London: Royal Aeronautical Society.

Reason, J.T. (1978). Motion sickness adaptation: A neural mismatch model. Journal of the Royal Society of Medicine, 71, 819-829.

Thomas, M.L., Reining, G., & Kelly, G. (1991). The Display for Advanced Research and Training: An "inexpensive" answer to tactical simulation. In H.M. Assenheim, R.A. Flasck, T.M. Lippert, & J. Bentz (Eds.), Large-Screen-Projection, Avionic, and Helmet-Mounted Displays Conference Proceedings. Bellingham, WA: SPIE-International Society for Optical Engineering.

Thomley, K.E., Kennedy, R.S., & Bittner, A.C., Jr. (1986). Development of postural equilibrium tests for examining environmental effects. Perceptual and Motor Skills, 63, 555-564.

Ungs, T.J. (1988). Simulator induced syndrome in Coast Guard aviators. Aviation, Space, and Environmental Medicine, 59, 267-272.

Ungs, T.J. (1989). Simulator induced syndrome: Evidence for long-term aftereffects. Aviation, Space, and Environmental Medicine, 60, 252-255.

APPENDIX A

IN-FLIGHT DISPLAY-RELATED QUESTIONS

A. Display-related questions for single-ship flights

1. Does your height above ground appear realistic?

Pilot response	IPs DART	IPs LFOVD	Mixed pilots DART	Mixed pilots LFOVD
	(N=15)	(N=16)	(N=7)	(N=7)
Yes	93.33	81.25	71.43	71.43
Not in relation to the trees, which are too tall	6.67	12.50	-	-
Appear to be lower than the actual height	-	6.25	28.57	28.57

Note. Each table value represents the percentage of pilots who gave the response.

2. Does the distance of objects appear realistic?

Pilot response	IPs DART	IPs LFOVD	Mixed pilots DART	Mixed pilots LFOVD
	(N=15)	(N=15)	(N=7)	(N=6)
Yes	86.67	100.	71.43	100.
Vehicles appear too small and farther than they actually are	6.67	-	14.29	-
Trees "pop" into the scene, which is unrealistic	6.67	-	14.29	-

Note: Each table value represents the percentage of pilots who gave the response.

3. Do the objects appear realistic?

Pilot response	IPs DART	IPs LFOVD	Mixed pilots DART	Mixed pilots LFOVD
	(N=15)	(N=15)	(N=6)	(N=6)
Yes	66.67	66.67	66.67	83.33
Trees and houses do; not the fields, lakes, cities, or mountain texture	20.00	20.00	-	16.67
Yes, in the AOI display	-	6.67	-	-
Some objects appear fuzzy (out of focus)	-	6.67	-	-
Yes, except the trees "pop" into the scene	-	-	33.33	-
Yes, but the road is hard to see	6.67	-	-	-
Ground and mountains appear out of focus	6.67	-	-	-

Note. Each table value represents the percentage of pilots who gave the response.

B. Display-related questions for formation flights

1. Do the lead aircraft features appear realistic?

Pilot response	Ips DART	Ips LFOVD	Mixed pilots DART	Mixed pilots LFOVD
	(N=16)	(N=16)	(N=5)	(N=7)
Yes	87.50	75.00	100.	85.71
Yes, in the AOI	-	6.25	-	-
No, too fuzzy (out of focus)	6.25	-	-	-
Yes, but slightly fuzzy	6.25	12.50	-	-
Yes, but vapor trails occasionally would be more realistic	-	6.25	-	-
Yes, but cannot see details as far as in actual flight	-	-	-	14.29

Note. Each table value represents the percentage of pilots who gave the response.

2. Do the lead aircraft dynamics appear realistic?

Pilot response	IPs DART	IPs LFOVD	Mixed pilots DART	Mixed pilots LFOVD
	(N=15)	(N=16)	(N=5)	(N=7)
Yes	80.00	100.	100.	100.
Some maneuvers are unrealistic	13.33	-	-	-
Roll rates should be faster	6.67	-	-	-

Note. Each table value represents the percentage of pilots who gave the response.

3. Does the aircraft separation appear realistic?

Pilot response	IPs DART	IPs LFOVD	Mixed pilots DART	Mixed pilots LFOVD
	(N=15)	(N=16)	(N=5)	(N=7)
Yes	80.00	100.	80.00	100.
Beyond 2,000 feet it appears farther than it actually is	13.33		20.00	-
It often appears closer than it actually is	6.67		-	-

Note. Each table value represents the percentage of pilots who gave the response.

APPENDIX B

PILOT RATINGS OF VISUAL SYSTEM CHARACTERISTICS

The DART and LFOVD display characteristics the pilots rated at the end of the simulator sessions are presented in the subsequent pages along with the means and medians of the ratings. The response alternatives were converted to the following scale values for the computation of the means and medians:

5 = Very Acceptable
4 = Acceptable
3 = Borderline
2 = Unacceptable
1 = Very Unacceptable

Table 1. Means and Medians of the Pilot Ratings for the DART Display Characteristics

Question number	Display characteristic	IPs (N=15)		Mixed pilots (N=5)	
		Mean	Median	Mean	Median
1	Horizontal field of view	4.47	5	4.40	4
2	Vertical field of view	4.47	5	4.40	4
3	Display brightness	4.47	5	4.20	4
4	Brightness uniformity within each window	4.60	5	4.60	5
5	Brightness uniformity from window to window	4.67	5	4.80	5
6	Brightness range from light to dark	4.40	5	4.20	4
7	Display resolution	4.00	4	3.80	4
8	Image movement lag to control inputs	4.27	4	4.00	4
9	Object-to-background contrast	4.73	5	4.00	4
10	Color uniformity within each window	4.67	5	4.40	5
11	Color uniformity from window to window	4.67	5	4.80	5
12	Color separation between adjacent objects and/or surfaces	4.33	4	4.40	4
13	Color fringes around objects or surfaces	4.27	4	4.20	4
14	Range of colors	4.13	4	4.20	4
15	Color saturation	4.40	4	4.40	4
16	Image transition from window to window	4.20	4	4.60	5
17	Simulated lead aircraft detail and dynamics	4.27	4	4.20	4

Table 1. (continued)

Question number	Display characteristic	IPs (N=15)		Mixed pilots (N=5)	
		Mean	Median	Mean	Median
18	Image obscuration by display window spacers	4.07	4	4.40	4
19	Attitude cuing from visual scene	4.53	5	4.40	4
20	Altitude cuing from visual scene	4.20	4	4.20	4
21	Depth cuing from visual scene	4.13	4	3.80	4
22	Motion cuing from visual scene	4.60	5	4.40	4
23	Scene alignment between adjacent windows	4.60	5	4.60	5
24	Viewing distance to windows	4.27	4	4.20	4
25	Image distortion within each window	4.53	5	4.20	4
26	Image distortion over all windows	4.60	5	4.20	4
27	Display noise (visible artifacts)	4.67	5	4.20	4
28	Object smearing	4.53	5	4.20	4
29	Display ghosting (double images)	4.80	5	4.40	4
30	Display flicker	4.53	5	4.20	4
31	Image scintillation (shimmering, sparkling)	4.73	5	4.20	4
32	Raster line visibility	4.67	5	4.40	4
33	Viewing space (area that head can be moved without disrupting visibility of image)	4.53	5	4.60	5
34	Edge continuity of objects and surfaces	4.40	5	4.60	5

Table 1. (concluded)

Question number	Display characteristic	IPs (N=15) Mean	IPs (N=15) Median	Mixed pilots (N=5) Mean	Mixed pilots (N=5) Median
35	Peripheral window "on-off" changes during head movements	4.40	4	4.20	4
36	HUD symbology brightness, dynamics, and location	4.40	5	4.60	5
37	Image jitter	4.60	5	4.20	4
38	Display aberrations (glare, reflections, scratches)	4.80	5	4.60	5
39	Scene content (trees, buildings, lakes, roads)	4.27	4	4.40	4
40	Image fidelity (realism)	4.33	4	4.00	4

Table 2. Means and Medians of the Pilot Ratings for the LFOVD AOI Display Characteristics

Question number	AOI display characteristic	IPs (N=15)		Mixed pilots (N=5)	
		Mean	Median	Mean	Median
1	Horizontal field of view for single-ship task	3.87	4	4.00	4.0
2	Vertical field of view for single-ship task	3.60	4	4.17	4.0
3	Horizontal field of view for formation flight task	3.80	4	3.83	4.0
4	Vertical field of view for formation flight task	3.40	4	3.67	4.0
5	Display resolution	4.07	4	4.33	4.0
6	Display brightness	4.40	5	4.33	4.0
7	Brightness uniformity across the display	4.47	5	4.50	4.5
8	Brightness range from light to dark	4.40	4	4.17	4.0
9	Object-to-background contrast	4.27	4	4.33	4.0
10	Color uniformity across the display	4.60	5	4.67	5.0
11	Color separation between adjacent objects and/or surfaces	4.60	5	4.50	4.5
12	Color fringes around objects or surfaces	4.60	5	4.33	4.0
13	Range of colors	4.27	4	4.17	4.0
14	Color saturation	4.40	4	4.33	4.0
15	Attitude cuing from visual scene	4.33	4	4.67	5.0
16	Altitude cuing from visual scene	4.47	5	4.33	4.0

Table 2. (concluded)

Question number	AOI display characteristic	IPs (N=15) Mean	Median	Mixed pilots (N=5) Mean	Median
17	Depth cuing from visual scene	4.47	5	4.33	4.0
18	Motion cuing from visual scene	4.53	5	4.67	5.0
19	Image distortion	4.53	5	4.50	4.5
20	Display noise (visible artifacts)	4.53	5	4.50	4.5
21	Object smearing	4.53	5	4.17	4.0
22	Display ghosting (double images)	4.80	5	4.67	5.0
23	Display flicker	4.20	4	4.67	5.0
24	Image scintillation (shimmering, sparkling)	4.73	5	4.67	5.0
25	Raster line visibility	4.60	5	4.33	4.0
26	Edge continuity of objects and surfaces	4.53	5	4.67	5.0

Table 3. Means and Medians of the Pilot Ratings for the LFOVD Background Display Characteristics

Question number	Background display characteristic	IPs (N=15)		Mixed pilots (N=5)	
		Mean	Median	Mean	Median
1	Horizontal field of view for single-ship task	3.90	4	3.83	4.0
2	Vertical field of view for single-ship task	3.73	4	4.00	4.0
3	Horizontal field of view for formation flight task	3.80	4	3.83	4.0
4	Vertical field of view for formation flight task	3.53	4	3.67	4.0
5	Display resolution	3.33	3	3.67	3.5
6	Display brightness	4.27	4	4.00	4.0
7	Brightness uniformity across the display	4.47	4	4.00	4.0
8	Brightness range from light to dark	4.27	4	4.17	4.0
9	Object-to-background contrast	4.00	4	4.00	4.0
10	Color uniformity across the display	4.40	4	4.33	4.0
11	Color separation between adjacent objects and/or surfaces	4.46	5	4.17	4.0
12	Color fringes around objects or surfaces	4.40	4	4.00	4.0
13	Range of colors	4.20	4	4.00	4.0
14	Color saturation	4.13	4	3.83	4.0
15	Attitude cuing from visual scene	4.20	4	4.67	5.0
16	Altitude cuing from visual scene	4.06	4	4.17	4.5
17	Depth cuing from visual scene	4.07	4	3.83	4.0

Table 3. (concluded)

Question number	Background display characteristic	IPs (N=15) Mean	IPs (N=15) Median	Mixed pilots (N=5) Mean	Mixed pilots (N=5) Median
18	Motion cuing from visual scene	4.40	4	4.50	4.5
19	Image distortion	4.46	4	4.00	4.0
20	Display noise (visible artifacts)	4.40	4	4.33	4.0
21	Object smearing	4.20	4	3.83	4.0
22	Display ghosting (double images)	4.73	5	4.67	5.0
23	Display flicker	4.27	4	4.67	5.0
24	Image scintillation (shimmering, sparkling)	4.67	5	4.50	4.5
25	Raster line visibility	4.60	5	4.00	4.0
26	Edge continuity of objects and surfaces	4.33	4	4.50	4.5

Table 4. Means and Medians of the Pilot Ratings for the General LFOVD Display Characteristics

Question number	General display characteristic	IPs (N=15) Mean	IPs (N=15) Median	Mixed pilots (N=5) Mean	Mixed pilots (N=5) Median
1	Image transition in and out of the AOI	3.93	4	3.67	3.5
2	Simulated lead aircraft detail and dynamics	4.60	5	4.50	4.5
3	Viewing distance to dome	4.73	5	4.67	5.0
4	Viewing space (area that head can be moved without disrupting visibility of image)	4.07	4	4.00	4.0
5	Blend area between AOI and background	3.73	4	3.83	4.0
6	Horizontal AOI excursion limits	3.67	4	3.83	4.0
7	Vertical AOI excursion limits	3.20	3	3.33	3.0
8	AOI lag to head movements	4.27	4	4.33	4.0
9	Image jitter	4.53	5	4.17	4.0
10	AOI tracking during head movements	4.13	4	4.17	4.0
11	Image movement lag to control inputs	4.40	4	4.50	4.5
12	Image alignment between AOI and background	4.47	4	4.33	4.0
13	Display aberrations (glare, reflections, scratches)	4.47	5	4.50	4.5
14	Scene content (trees, building, lakes, roads)	4.47	5	4.33	4.0
15	Head movements induced by AOI	3.80	4	3.83	4.0
16	Image fidelity (realism)	4.40	4	4.00	4.0

APPENDIX C

MEANS OF THE SIGNIFICANT MAIN EFFECTS AND INTERACTIONS FOR THE SINGLE-SHIP FLIGHT PERFORMANCE MEASURES

Table 1. Means of the Significant Main Effects for the Single-Ship Flight Performance Measure: Mean G-Force

Main effect	Mean
Visual system	
DART	2.44
LFOVD	2.32
Flight time (min)	
5	2.20
10	2.20
15	2.21
20	2.90

Table 2. Means of the Significant Main Effects for the Single-Ship Flight Performance Measure: Std. Dev. G-Force

Main effect	Mean
Visual system	
DART	2.60
LFOVD	1.87
Flight time (min)	
5	2.18
10	2.13
15	2.11
20	2.51

Table 3. Means of the Significant Main Effects and Interactions for the Single-Ship Flight Performance Measure: Mean Altitude

Main effect	Mean
Visual system	
DART	398.44
LFOVD	347.25
Flight time (min)	
5	403.46
10	362.68
15	321.55
20	403.69

Interaction	Means
Visual system x Pilot group	

	Visual system	
Pilot group	DART	LFOVD
IPs	362.79	339.51
Mixed	505.40	370.47

Flight time x Pilot group

	Flight time (min)			
Pilot group	5	10	15	20
IPs	373.07	331.25	304.05	396.22
Mixed	494.62	456.95	374.06	426.10

Table 4. Means of the Significant Main Effect and Interaction for the Single-Ship Flight Performance Measure: Std. Dev. Altitude

Main effect	Mean
Flight time (min)	
5	209.94
10	197.37
15	171.59
20	228.68

Interaction	Means
Visual system x Pilot group	

	Visual system	
Pilot group	DART	LFOVD
IPs	172.47	221.54
Mixed	243.06	190.06

Table 5. Means of the Significant Main Effect and Interaction for the Single-Ship Flight Performance Measure: Mean Airspeed

Main effect	Mean
Flight time (min)	
5	453.44
10	452.78
15	450.52
20	434.38

Interaction	Means
Visual system x Flight time x Pilot group	

Visual system - Pilot group	Flight time (min)			
	5	10	15	20
DART - IPs	448.68	446.87	449.10	437.29
DART - Mixed	473.32	456.22	448.22	429.14
LFOVD - IPs	452.95	452.98	449.01	432.87
LFOVD - Mixed	449.28	466.46	461.66	435.38

Table 6. Means of the Significant Main Effects and Interactions for the Single-Ship Flight Performance Measure: Std. Dev. Airspeed

Main effect	Mean
Visual system	
DART	33.40
LFOVD	37.29
Flight time (min)	
5	33.60
10	36.21
15	27.57
20	44.01

Interaction	Means

Flight time x Pilot group

Pilot group	Flight time (min) 5	10	15	20
IPs	30.00	36.64	28.48	42.20
Mixed	44.42	34.93	24.82	49.41

Visual system x Flight time x Pilot group

Visual system - Pilot group	Flight time (min) 5	10	15	20
DART - IPs	26.75	34.22	29.90	42.06
DART - Mixed	48.28	30.80	19.11	37.43
LFOVD - IPs	33.24	39.07	27.06	42.35
LFOVD - Mixed	40.56	39.06	30.53	61.39

Table 7. **Means of the Significant Main Effect and Interaction for the Single-Ship Flight Performance Measure: Mean Absolute Pitch Angle**

Main effect	Mean
Flight time (min)	
5	3.96
10	3.47
15	3.22
20	4.48

Interaction	Means
Visual system x Pilot group	

Pilot group	Visual system DART	Visual system LFOVD
IPs	3.69	3.91
Mixed	4.01	3.43

Table 8. Means of the Significant Main Effect and Interaction for the Single-Ship Flight Performance Measure: Std. Dev. Pitch Angle

Main effect	Mean
Flight time (min)	
5	5.09
10	4.46
15	4.20
20	6.15

Interaction	Means
Visual system x Pilot group	

Pilot group	Visual system DART	Visual system LFOVD
IPs	4.89	5.18
Mixed	5.34	4.25

Table 9. Means of the Significant Main Effect for the Single-Ship Flight Performance Measure: Mean Absolute Bank Angle

Main effect	Mean
Flight time (min)	
5	44.81
10	45.39
15	45.14
20	54.02

Table 10. Means of the Significant Main Effects for the Single-Ship Flight Performance Measure: Std. Dev. Bank Angle

Main effect	Mean
Visual system	
DART	56.87
LFOVD	55.55
Flight time (min)	
5	54.22
10	54.09
15	53.84
20	62.69

Table 11. Means of the Significant Main Effects for the Single-Ship Performance Measure: Mean Absolute Pitch Rate

Main effect	Mean
Visual system	
DART	7.20
LFOVD	5.38
Flight time (min)	
5	5.86
10	5.84
15	5.64
20	7.83

Table 12. Means of the Significant Main Effects for the Single-Ship Performance Measure: Std. Dev. Pitch Rate

Main effect	Mean
Visual system	
DART	8.75
LFOVD	6.08
Flight time (min)	
5	7.16
10	7.11
15	6.97
20	8.43

Table 13. **Means of the Significant Main Effects and Interaction for the Single-Ship Performance Measure: Mean Absolute Roll Rate**

Main effect	Mean
Visual system	
DART	15.98
LFOVD	12.42
Flight time (min)	
5	13.44
10	13.50
15	13.32
20	16.55

Interaction	Means
Flight time x Pilot group	

Pilot group	Flight time (min) 5	10	15	20
IPs	13.98	13.80	13.19	17.14
Mixed	11.82	12.60	13.71	14.78

Table 14. Means of the Significant Main Effects for the Single-Ship Performance Measure: Std. Dev. Roll Rate

Main effect	Mean
Visual system	
DART	32.67
LFOVD	21.80
Flight time (min)	
5	25.90
10	25.81
15	26.09
20	31.13

Table 15. Means of the Significant Main Effect for the Single-Ship Performance Measure: Mean Absolute Yaw Rate

Main effect	Mean
Flight time (min)	
5	1.88
10	1.93
15	1.89
20	2.43

Table 16. Means of the Significant Main Effects for the Single-Ship Performance Measure: Std. Dev. Yaw Rate

Main effect	Mean
Visual system	
DART	3.26
LFOVD	2.43
Flight time (min)	
5	2.63
10	2.66
15	2.63
20	3.47

APPENDIX D

MEANS OF THE SIGNIFICANT MAIN EFFECTS AND INTERACTIONS FOR THE FORMATION FLIGHT PERFORMANCE MEASURES

Table 1. Means of the Significant Main Effect for the Formation Flight Performance Measure: Mean G-Force

Main effect	Mean
Flight time (min)	
5	2.66
10	2.55
15	3.12
20	2.85

Table 2. Means of the Significant Main Effects and Interaction for the Formation Flight Performance Measure: Std. Dev. G-Force

Main effect	Mean
Visual system	
DART	3.12
LFOVD	2.28
Flight time (min)	
5	2.66
10	2.68
15	2.81
20	2.59

Interaction	Means
Visual system x Flight time	

	Flight time (min)			
Visual system	5	10	15	20
DART	3.20	3.18	3.23	2.87
LFOVD	2.15	2.21	2.41	2.34

Table 3. Means of the Significant Main Effect for the Formation Flight Performance Measure: Mean Altitude

Main effect	Mean
Flight time (min)	
5	940.14
10	906.59
15	833.21
20	723.86

Table 4. Means of the Significant Main Effect for the Formation Flight Performance Measure: Std. Dev. Altitude

Main effect	Mean
Flight time (min)	
5	686.59
10	666.82
15	539.80
20	482.16

Table 5. **Means of the Significant Main Effects for the Formation Flight Performance Measure: Mean Airspeed**

Main effect	Mean
Visual system	
DART	498.43
LFOVD	515.03
Flight time (min)	
5	524.21
10	535.97
15	502.68
20	464.02

Table 6. Means of the Significant Main Effect for the Formation Flight Performance Measure: Std. Dev. Airspeed

Main effect	Mean
Flight time (min)	
5	39.96
10	41.11
15	57.71
20	68.12

Table 7. Means of the Significant Main Effect for the Formation Flight Performance Measure: Mean Absolute Bank Angle

Main effect	Mean
Flight time (min)	
5	52.24
10	56.10
15	56.96
20	53.13

Table 8. Means of the Significant Main Effect for the Formation Flight Performance Measure: Std. Dev. Bank Angle

Main effect	Mean
Flight time (min)	
5	63.29
10	66.93
15	66.70
20	63.36

Table 9. Means of the Significant Main Effects and Interaction for the Formation Flight Performance Measure: Mean Absolute Pitch Rate

Main effect	Mean
Visual system	
DART	7.99
LFOVD	5.79
Flight time (min)	
5	6.42
10	6.22
15	7.49
20	7.31

Interaction	Means
Visual system x Flight time	

Visual system	Flight time (min) 5	10	15	20
DART	7.84	7.49	8.60	8.05
LFOVD	5.08	5.01	6.44	6.64

Table 10. Means of the Significant Main Effects and Interaction for the Formation Flight Performance Measure: Std. Dev. Pitch Rate

Main effect	Mean
Visual system	
DART	9.18
LFOVD	6.00
Flight time (min)	
5	7.28
10	7.06
15	7.88
20	7.95

Interaction	Means
Visual system x Flight time	

Visual system	Flight time (min) 5	10	15	20
DART	9.24	8.87	9.44	9.18
LFOVD	5.41	5.33	6.39	6.84

Table 11. Means of the Significant Main Effects for the Formation Flight Performance Measure: Mean Absolute Roll Rate

Main effect	Mean
Visual system	
DART	16.22
LFOVD	13.63
Flight time (min)	
5	13.80
10	14.59
15	15.83
20	15.32

Table 12. Means of the Significant Main Effects for the Formation Flight Performance Measure: Std. Dev. Roll Rate

Main effect	Mean
Visual system	
DART	33.32
LFOVD	23.80
Flight time (min)	
5	26.71
10	27.68
15	29.89
20	29.39

Table 13. Means of the Significant Main Effects for the Formation Flight Performance Measure: Mean Absolute Yaw Rate

Main effect	Mean
Visual system	
DART	2.02
LFOVD	1.76
Flight time (min)	
5	1.71
10	1.70
15	2.03
20	2.12

Table 14. Means of the Significant Main Effects and Interaction for the Formation Flight Performance Measure: Std. Dev. Yaw Rate

Main effect	Mean
Visual system	
DART	3.02
LFOVD	2.27
Flight time (min)	
5	2.38
10	2.67
15	2.84
20	3.05

Interaction	Means
Visual system x Flight time	

Visual system	Flight time (min) 5	10	15	20
DART	2.82	2.54	3.30	3.47
LFOVD	1.97	2.02	2.40	2.68

Table 15. Means of the Significant Interaction for the Formation Flight Performance Measure: Mean Slant Range

Interaction	Means
Visual system x Pilot group	

Pilot group	Visual system DART	Visual system LFOVD
IPs	1453.12	1562.50
Mixed	2162.55	1649.07

Table 16. Means of the Significant Interactions for the Formation Flight Performance Measure: Std. Dev. Slant Range

Interaction	Means
Visual system x Flight time	

Visual system	Flight time (min) 5	10	15	20
DART	1613.47	744.82	951.40	1059.51
LFOVD	684.82	675.32	1053.20	1098.86

Visual system x Pilot group

Pilot group	Visual system DART	LFOVD
IPs	860.15	900.42
Mixed	1732.34	803.29

APPENDIX E

GUIDELINES FOR ALLEVIATING SIMULATOR SICKNESS WITH THE DART AND LFOVD VISUAL SIMULATION SYSTEMS

Guidelines for Alleviating Simulator Sickness with the DART and LFOVD Visual Simulation Systems

---

Flight Monitor (instructor/researcher) Responsibilities

1. Determine if the subjects have previously experienced sickness in a simulator.
2. Learn the signs and symptoms of simulator sickness.
3. Be aware of the factors that may contribute to a subject's susceptibility:
    - a. medication
    - b. sleep loss
    - c. flu
    - d. respiratory illness
    - e. head cold
    - f. ear infection
    - g. ear blocks
    - h. upset stomach
    - i. hangover
    - j. emotional stress
    - k. menstruation
4. Provide a briefing to the subjects that simulator sickness can occur and that it is not an abnormal reaction.
5. Observe the subjects before, during, and after the simulation.
6. If sickness occurs during the simulator session, the subject should not return until all symptoms have subsided, usually 10 to 12 hours.
7. Make sure the subjects are over the symptoms before they leave the facility.
8. If the symptoms are severe, have the subject report to the flight surgeon.

User-Oriented Countermeasures

1. Minimize head movements and keep head position within the design eye area.
2. If the subject recently experienced simulator sickness:
    - a. limit the initial duration and variety of the flights,
    - b. turn off one or more visual channels initially,
    - c. provide night missions before day missions,
    - d. if eyestrain previously occurred, schedule morning flights, and
    - e. if "fullness of the head" or a persistent headache previously occurred, schedule afternoon flights.

Procedural Countermeasures

1. Start with simple maneuvers and gradually increase in intensity.
2. Minimize aircraft maneuvering near the ground at first.
3. Minimize rapid changes in altitude, abrupt rolls, and porpoising during the initial simulator exposures.

Guidelines for Alleviating Simulator Sickness in the DART and LFOVD Visual Simulation Systems (concluded)

4. Avoid freezing the simulator in the initial sessions. If not possible, freeze only after recovery to straight-and-level flight. Never freeze the simulator in unusual attitudes.
5. Do not reset or rapidly slew the aircraft when the scene is visible; blank the scene first.
6. During breaks and before entering or leaving the simulator, blank the scene. If this is not possible, position the simulated aircraft at 0 degrees pitch, yaw, and roll.
7. Ask the subjects to look down or close their eyes if the scene cannot be blanked when the simulator is repositioned.
8. Limit exposure duration and provide frequent breaks.
9. Simulator flights should not be scheduled on the same day as aircraft flights.

Engineering Countermeasures

1. Align the image projectors and display screens.
2. Avoid lags between control inputs and image movement.
3. Correct display anomalies, such as:
   a. color imbalances within and between viewing windows,
   b. misalignment of the horizon,
   c. display flicker, and
   d. conflicting information between the flight instruments and the visual scene.
4. Maintain a comfortable cockpit temperature.
5. Provide proper airflow.
6. Calibrate and maintain the simulator.